NASA TN D-4109

# A CRITICAL ANALYSIS OF VARIOUS NUMERICAL INTEGRATION METHODS FOR COMPUTING THE FLOW OF A GAS IN CHEMICAL NONEQUILIBRIUM

By Harvard Lomax and Harry E. Bailey

Ames Research Center
Moffett Field, Calif.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – CFSTI price $3.00

# A CRITICAL ANALYSIS OF VARIOUS NUMERICAL INTEGRATION METHODS FOR COMPUTING THE FLOW OF A GAS IN CHEMICAL NONEQUILIBRIUM

By Harvard Lomax and Harry E. Bailey

Ames Research Center

## SUMMARY

The efficiency and accuracy of several different techniques proposed for the numerical integration of ordinary differential equations with widely varying damping properties are examined. The methods analyzed include the standard Runge-Kutta method, a modified Runge-Kutta method proposed by Treanor, predictor-corrector schemes, and implicit procedures. First, the methods are studied as they apply to linear coupled differential equations with constant coefficients. The conclusions drawn from these studies are then tested in nonlinear cases by numerical calculations made for a gas in chemical nonequilibrium behind a normal shock wave. An implicit method is strongly recommended if the local eigenvalues are negative (damping) and widely separated.

## INTRODUCTION

The numerical integration of the nonlinear equations arising in the study of a gas flowing in chemical nonequilibrium poses, in certain cases, a severe numerical stability problem. The existence of such a problem is well known (see refs. 1-5). It is also well known that numerical solutions involving the use of standard Runge-Kutta and predictor-corrector methods (referred to as conventional methods) can be prohibitively expensive because they require excessive machine running time.

Several attempts have been made to overcome the difficulty by introducing numerical methods specifically designed to cope with the problem. These methods fall into two principal categories. In one, the nonlinear differential equations are reduced to nonlinear difference equations, as in the conventional methods, but the coefficients in the differencing equations contain certain parameters which depend on the solution as it proceeds. If the differential equations were uncoupled, these parameters would be the local eigenvalues of the individual equations. A typical example of this class is the method proposed by Treanor. In the other category, the differential equations are first locally linearized and the resulting linearized form is solved either exactly (ref. 6), approximately (ref. 7), or by finite difference methods (ref. 8).

A basic objective which motivated the numerical research reported herein was to compare the efficiency of the various numerical methods available. To accomplish such an objective, it is essential that we know the fundamental relationship of a set of coupled differential equations and a corresponding set of coupled difference equations obtained from the former by the application of any given differencing scheme. To find such a relationship that applies rigorously to general nonlinear equations is quite outside the scope of this paper. However, the fundamental relations regarding the stability and accuracy of numerical solutions to coupled linear differential equations with constant coefficients can be rigorously defined. As a result of these definitions the numerical methods mentioned can be compared objectively as they apply to such linear equations. We then hypothesize that this comparison can be used to estimate their usefulness in nonlinear cases.

## LIST OF IMPORTANT SYMBOLS

| | |
|---|---|
| $[\ ]$ | matrix of enclosed quantity |
| $[\ ]^{-1}$ | inverse of matrix |
| $\overline{A}$ | area |
| $[A_n]$ | matrix in locally linearized equations (see eq. (4)) |
| det( ) | determinant of enclosed quantity |
| $\vec{F}$ | derivative of $\vec{w}$ with respect to s |
| $\vec{F}_n$ | value of $\vec{F}$ at step n |
| H | effective distance a numerical method advances the integration after time for two evaluations of the derivatives |
| $h_j$ | enthalpy of the jth species |
| [I] | unit matrix |
| $K_i$ | equilibrium constant for ith reaction |
| m | number of equations |
| N | number of chemical species |
| P | Treanor parameter (see eq. (10)) |
| $Q^j$ | production of species j in moles per unit volume per unit time |
| R | universal gas constant |

| | |
|---|---|
| $s$ | independent variable |
| $T$ | temperature |
| $u$ | velocity |
| $\vec{v}$ | dependent variable in uncoupled equations (see eq. (5)) |
| $\vec{w}$ | dependent variable in coupled equations (see eq. (2)) |
| $\vec{w}'$ | derivative of $\vec{w}$ with respect to $s$ |
| $x$ | distance downstream of normal shock |
| $\gamma_j$ | molar concentration of $j$th species, moles per unit mass |
| $\epsilon$ | upper bound of error permitted in numerical solution |
| $\lambda_1$ | parasitic eigenvalue |
| $\lambda_2$ | driving eigenvalue |
| $\rho$ | density |
| $\chi_i$ | degree of nonequilibrium of $i$th reaction |

Superscripts

| | |
|---|---|
| $\rightarrow$ | vector |
| $T$ | transpose of vector |

## THE BASIC NONLINEAR EQUATIONS

The equations governing inviscid, one-dimensional, nonequilibrium flow can be written

$$\left.\begin{aligned}
&\rho\bar{A}\frac{du}{dx} + u\bar{A}\frac{d\rho}{dx} = -\rho u\bar{A}_x \\
&\rho u\frac{du}{dx} + \left(RT\sum_1^N \gamma_i\right)\frac{d\rho}{dx} + \left(\rho R\sum_1^N \gamma_i\right)\frac{dT}{dx} + \rho RT\sum_1^N \frac{d\gamma_i}{dx} = 0 \\
&u\frac{du}{dx} + \left(\sum_1^N \gamma_i c_{p_i}\right)\frac{dT}{dx} + \sum_1^N h_i\frac{d\gamma_i}{dx} = 0 \\
&\rho u\frac{d\gamma_1}{dx} = Q^1(\rho,T,\gamma_1,\ldots,\gamma_N) \\
&\quad\vdots \\
&\rho u\frac{d\gamma_N}{dx} = Q^N(\rho,T,\gamma_1,\ldots,\gamma_N)
\end{aligned}\right\} \quad (1)$$

Although the equations are written for general one-dimensional channel flows, only the results for flow behind a normal shock, for which $\bar{A} = 1$ and $\bar{A}_x = 0$, are presented. This is done deliberately to isolate the basic <u>numerical</u> problems discussed in the following sections. The fluid dynamic equations were differenced and numerically integrated just like the chemical equations even though they could have been integrated analytically. Thus, all the dependent variables are treated alike as they would be in a general two-dimensional flow.

Equations (1) can be written in matrix-vector form by introducing the vector $\vec{w}^*$, such that its transpose $\vec{w}^{*T} = (u,\rho,T,\gamma_1,\ldots,\gamma_N)$, the vector $\vec{c}^*$, such that $\vec{c}^{*T} = (-\rho u\bar{A}_x,0,0,Q^1,\ldots,Q^N)$, and the matrix $[B^*]$. The result is

$$[B^*]\frac{d\vec{w}^*}{dx} = c^*$$

In the general case the matrix $[B^*]$ can become singular (this occurs when $u$ equals the sonic velocity). Some of the numerical difficulties brought about by this occurrence can be avoided by introducing the new independent variable $s$ such that

$$x' = \frac{dx}{ds} = \det(B^*)$$

and inverting the equation and multiplying both sides by $\det(B^*)$ to obtain

$$\vec{w}^{*\prime} \equiv \frac{d\vec{w}^*}{ds} = \vec{F}^* \equiv \det(B^*)[B^*]^{-1}\vec{c}^*$$

where $\det(B^*)[B^*]^{-1}$ is the adjoint of $[B^*]$. Finally, we define the new vectors $\vec{w}$ and $\vec{F}$ with one element more than their starred counterparts

$$\vec{w}^T = (\vec{w}^{*T}, x)$$

or

$$\vec{w}^T = (u, \rho, T, \gamma_1, \ldots, \gamma_N, x)$$

and

$$\vec{F}^T = [\vec{F}^{*T}, \det(B^*)]$$

This provides a set of $m = N + 4$ simultaneous equations

$$\frac{d\vec{w}}{ds} = \vec{F}(\vec{w}) \tag{2}$$

in which there is no explicit dependence of $\vec{F}$ on the independent variable s. These are the basic nonlinear equations that motivated the numerical studies reported herein.

## LOCAL LINEARIZATION AND PARASITIC EIGENVALUES

Consider the set of equations

$$\frac{d\vec{w}}{ds} = \vec{w}^{\prime} = \vec{F}(\vec{w}) \tag{3}$$

These equations are autonomous; that is, each equation is of the form

$$w_i^{\prime} = F_i(w_1, w_2, \ldots, w_m)$$

where $F_i$ has no explicit dependence on the independent variable s. If each $F_i$ is expanded about a local point referenced as n, where $s = nh$,

$$w_i^{\prime} = F_{in} + (w_1 - w_{1n})\left(\frac{\partial F_i}{\partial w_1}\right)_n + \ldots + (w_m - w_{mn})\left(\frac{\partial F_i}{\partial w_m}\right)_n + O[(\vec{w} - \vec{w}_n)^2]$$

Let the elements $(a_{ij})_n$ of a matrix $[A_n]$ be $(\partial F_i/\partial w_j)_n$. Note that the neglected higher derivative terms involve $(\vec{w} - \vec{w}_n)^2$. If $\vec{w} = \vec{w}_{n+1}$, this can be written $h^2[(\vec{w}_{n+1} - \vec{w}_n)/h]^2$ or $h^2(\vec{w}'_n)^2$ plus terms of $O(h^3)$. Thus we find

$$\vec{w}' = [A_n]\vec{w} + \vec{F}_n - [A_n]\vec{w}_n + O(h^2) \tag{4}$$

Equation (4) gives, neglecting higher order terms, a locally linearized form of the original equations. Further, since the original equations were autonomous, the linearized equations have constant coefficients. It is well known that the complementary solution of equation (4) can be written

$$w_i = \sum_{j=1}^{J} \left[ C_{ij1} + C_{ij2}s + \ldots + C_{ijK(j)}s^{K(j)-1} \right] e^{\lambda_j s}, \qquad i = 1, 2, \ldots, m$$

where $K(j)$ is the multiplicity of the $j$th root to the characteristic equation, $J$ is the number of distinct roots, $\lambda_j$, and $m$ is the number of equations. From a somewhat different point of view, the values of $\lambda_j$ (which can be complex) are equal in magnitude, sign, and multiplicity to the eigenvalues in the matrix $[A_n]$.

Now it is typical of equations for nonequilibrium flow that in certain regions some of the eigenvalues are large, negative, real numbers and some are relatively much smaller in magnitude. Let these be represented by $(\lambda_1)_n = -1000\mu$ and $(\lambda_2)_n = -\mu$, respectively. When integrating equations with such eigenvalues, two cases can occur. One, we wish to resolve the effect of $(\lambda_1)_n$ over the small region where $e^{-1000\mu s}$ is significant. Then we must perform our calculations at points spaced very close together. Two, $s$ is large enough for $e^{-1000\mu s}$ to be negligible compared to $e^{-\mu s}$. Then we need only use the much coarser spacing that resolves $e^{-\mu s}$. If the solution could be carried out analytically, this would pose no problem. However, if it is carried out numerically, using conventional Runge-Kutta or predictor-corrector schemes, violent instabilities occur for the coarse spacing. For this reason, when a set of differential equations has eigenvalues such as $(\lambda_1)_n$ and $(\lambda_2)_n$ in the above example, we refer to those like $(\lambda_1)_n$ as "parasitic" eigenvalues and those like $(\lambda_2)_n$ as "driving" eigenvalues. Sets of equations having this property are often referred to as "stiff" equations.

In order to discuss numerical methods used to solve equations with parasitic eigenvalues, we need the following brief discussion of some results from numerical analysis.

## A THEOREM REGARDING THE NUMERICAL SOLUTION OF ORDINARY DIFFERENTIAL EQUATIONS

Equation (4) represents a set of coupled, ordinary, differential equations with constant coefficients. In general, these can be uncoupled

insofar as they can be reduced to the Jordan canonical form (see ref. 9). For the purposes of numerical research it is often convenient to construct coupled equations from simple sets of uncoupled ones. Since this is the approach taken in appendix B, and since it illustrates the connection between coupled and uncoupled sets, we proceed to outline it here.

Consider the special uncoupled set of linear equations with constant coefficients

$$v_1' = \lambda_1 v_1 + f_1$$

$$v_2' = \lambda_2 v_2 + f_2$$

$$\cdot \qquad \cdot \quad \cdot$$

$$\cdot \qquad \cdot \quad \cdot$$

$$\cdot \qquad \cdot \quad \cdot$$

These are special in the sense that their solutions have no terms of the form $C_{ijk}s^k$ with $k > 0$ in the complementary solution of equation (4); in other words, no coupled set formed from them will have a matrix with multiple eigenvalues. They can be written

$$\vec{v}' = [L]\vec{v} + \vec{f} \tag{5}$$

where [L] is a diagonal matrix (i.e., all off-diagonal elements are zero). Let

$$\vec{w} = [B]\vec{v} \tag{6}$$

where [B] is an <u>arbitrary</u> nonsingular matrix. Then, if

$$[A] = [B][L][B]^{-1} \tag{7}$$

and

$$\vec{g} = [B]\vec{f}$$

we can write

$$\vec{w}' = [A]\vec{w} + \vec{g} \tag{8}$$

where the eigenvalues of [L] and [A] are <u>identical</u>. Introduce the linear operator $l_h$ to represent any conventional[1] Runge-Kutta or predictor-corrector method with step size h. It is essential that, in a given step,

[1]More precisely, any set of linear difference-differential equations with constant coefficients applied in the same way to all of the simultaneous differential equations. All references to conventional numerical methods in this report assume a choice from this set.

the identical method be applied to all the equations. Then, if the numerical integration is started at the step $n - j$, at the step $n$ we have

$$\vec{v}_n = \ell_h(\vec{v}_j, \vec{v}_j', \vec{f}_j) \quad , \qquad j = n, n-1, \ldots, n-j$$

The value of $\vec{v}_n$ differs from the analytical value of $\vec{v}$ by an amount equal to the error introduced by the numerical method.

Consider the following theorem whose proof is given in reference 10.

Theorem I: If $\vec{v}_n = \ell_h(\vec{v}_j, \vec{v}_j', \vec{f}_j)$ and $w_n = \ell_h(\vec{w}_j, \vec{w}_j, \vec{g}_j)$, then

$\vec{v}_n = [B]\vec{w}_n + er$, where er depends only on the round-off process used in the computation.

The theorem can be paraphrased as follows: except for round-off error, solutions are the same whether or not a set of linear differential equations with constant coefficients is first uncoupled and then integrated numerically, or first integrated numerically and then uncoupled, no matter what conventional numerical method is employed, provided the same method is used on all of the equations. Two corollaries follow, both of which assume that machine limitations such as round off and floating overflow can be neglected.

(1) Both the stability and accuracy of a conventional numerical method applied to equation (8) are independent of the magnitude and sign of the elements of [A] except as those elements determine the eigenvalues.

(2) A step size that is chosen to resolve the effects of driving eigenvalues can be used to integrate equation (8) and the numbers obtained from the solution contain, to the desired accuracy, all information concerning the effects of the driving eigenvalues regardless of errors brought about by the parasitic ones.

Corollary (2) is literally true even if the numerical method is unstable for the parasitic eigenvalues. However, the computing machine limitations neglected in formulating the corollary make this an impractical result for our purposes even if we were to uncouple the results. The important consequence of the corollaries is that to provide accurate and usable solutions to problems with parasitic eigenvalues, we need only provide a method that is stable for large, negative, real values of $\lambda_1 h$ and accurate for any complex $\lambda_2 h$ relatively much smaller in magnitude. We next discuss some methods from this point of view.

## APPRAISAL OF METHODS

With the preceding material as a background, we can now discuss the relative merits of several procedures used for the numerical integration of coupled equations with parasitic eigenvalues. These procedures include

explicit methods, implicit methods, and special methods that have been constructed for problems of this type. In order to compare such methods fairly, we first introduce a representative step size such that

$H \equiv$ the effective distance a numerical method advances the integration after a time equal to the time required for two evaluations of the derivatives.

Since in the numerical solution of nonequilibrium problems, the time required to calculate the derivatives is the predominant factor, H is the significant reference, rather than h, the step size used in the computation. Furthermore, we define the term real stability boundary as the largest real negative value of $\lambda H$ for which a numerical method is stable when used to integrate a set of coupled, linear, differential equations having $\lambda$ as an eigenvalue.

## Conventional Explicit Methods

These methods encompass all standard predictor-corrector and Runge-Kutta techniques. About the simplest of such techniques, and one that has an accuracy that is acceptable for many practical cases, is given by

$$\left.\begin{aligned} v_{n+1}^{(1)} &= v_n + hv_n' \\ v_{n+1} &= v_n + \frac{1}{2} h \left( v_{n+1}^{(1)\prime} + v_n' \right) \end{aligned}\right\} \tag{9}$$

This method is referred to either as an Euler predictor with a modified Euler corrector or as a second-order, Runge-Kutta method. Actually, the method is unstable for problems with high-frequency, low-amplitude noise (see ref. 10), but this does not occur in the problems under consideration since the parasitic eigenvalues are real, not imaginary. Equations (9) have a truncation error led by $(1/6)(\lambda H)^3$ and a real stability boundary equal to -2.0. Higher order Runge-Kutta methods are more accurate but less stable. For example, the third-order Runge-Kutta (Heune) method has a real stability boundary equal to about -1.6, and the standard fourth-order, Runge-Kutta method is limited still further to about -1.4. Most conventional predictor-corrector methods (Hamming's, Adams-Moulton, etc.) have real stability boundaries lying between -1 and 0.

For all conventional explicit methods the step size is dictated by the largest eigenvalue of the system no matter whether it is parasitic or driving.

## Treanor's Explicit Method

Treanor's method (ref. 3) is explicit and was designed specifically to cope with the numerical integration of stiff equations. In constructing this method it is assumed that the basic equations can be approximated by a linearized form, but not the form given in equation (4). Instead, each equation is

differenced as if it could be written[2]

$$v_i' = -(P_i)_n v_i + (A_i)_n + (B_i)_n s + (C_i)_n s^2 \qquad (10)$$

where $(P_1)_n$, $(P_2)_n$, . . . are constants in a given step, but are allowed to vary from equation to equation, and $A_i$, $B_i$, and $C_i$ are all functions of $P_i$ only. The parameters $(P_i)_n$ would be the local eigenvalues of the individual equations if the equations were uncoupled. In practice they are formed by ratioing certain terms in the first two calculations made at the intermediate steps in a standard, fourth-order, Runge-Kutta process.

Certain general statements can be made about this method. It reverts to the Runge-Kutta method if $h$ is made small enough. Of course, if this property is used, the method will be no improvement on the conventional one. The method is excellent if the differential equations are "nearly" uncoupled. Unfortunately, however, it is not simple to make an a priori estimation of the degree of coupling in sets of equations.

The most serious difficulty that can arise in the general use of Treanor's method is discussed next. First, one can show (see appendix A) that if, in a given step, different differencing schemes are used on different equations in a coupled differential set, theorem I no longer holds. It follows immediately that a method which permits $P_i$ in equation (10) to vary with $i$ will give solutions whose accuracy and stability will, in general, depend on the individual elements in $[A_n]$. Since, for a fixed set of eigenvalues, these elements can have an extreme variation, such a dependency can lead to serious numerical errors. The actual values that the parameters $(P_i)_n$ are likely to take when found by Treanor's method are studied in appendix B. It is shown that they are nearly always about the same and approximately equal in magnitude to the largest parasitic eigenvalue detectable in $w$. However, the fact that this is not always the case is the principal deterrent to the general recommendation of the method.

If the parameters $P_i$ in the differencing scheme proposed by Treanor are made to be the same for all $i$ in a given step, theorem I does apply to the results. In such a case the method, when used to integrate coupled, linear, differential equations with constant coefficients, can be subjected to a complete analysis. This has been carried out and is presented in appendix C. The results regarding the stability for real negative eigenvalues are presented in figure 1. Since theorem I applies, the method acts on each eigenvalue as if the others were not present and a single figure gives the whole stability picture. If all the eigenvalues are such that $-2 < \lambda_j h < 0$, the Treanor differencing scheme with fixed $P$ is stable for $0 < Ph < \infty$. On the other hand, if an eigenvalue exists such that for it $\lambda h < -2$, a stability corridor starts to form, and the choice of $P$ becomes critical. Two situations can arise. First, if there is only one parasitic eigenvalue, or if all

[2]Another method which assumes this kind of local approximation is given in reference 11. However, because of the ambiguity in just how the values of $(A_i)_n$, $(B_i)_n$, and $(C_i)_n$ are to be formed in the case $\vec{v}' = \vec{F}(\vec{v},s)$, this method cannot be analyzed in detail.

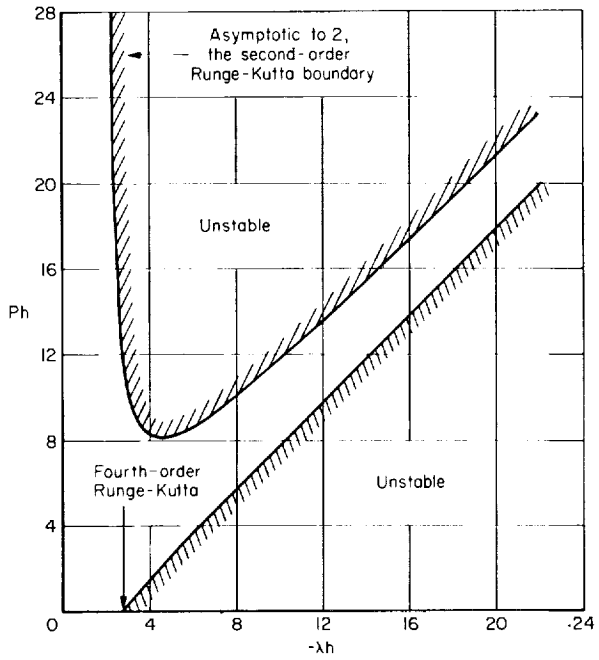



Figure 1.- Stability boundaries for Treanor's method for real negative $\lambda$.

the parasitic eigenvalues are clustered inside the corridor, the method is stable for indefinitely large separations between the parasitic and driving eigenvalues, provided P is appropriately chosen. Second, if two or more parasitic eigenvalues do not lie in the corridor, the method is still stable for coupled equations in which $0 > \lambda_j h > -10$ if Ph is 8. Since the method requires four evaluations of the derivatives in a single calculation step, this gives a real stability boundary $\lambda H$ equal to about 5, an improvement of about 3.6 over the standard, fourth-order, Runge-Kutta method. These observations suggest certain modifications to Treanor's method if it is to be used in general; two such modifications are discussed at the end of appendix B.

## Implicit Methods

The methods discussed above are explicit. It has been known for some time that certain implicit numerical methods are unconditionally stable for all real negative values of $\lambda h$. Probably the most widely used of these is the implicit modified Euler method described by

$$v_{n+1} = v_n + \frac{1}{2} h(v'_{n+1} + v'_n) \tag{11}$$

which is even stable for all complex values of $\lambda h$ with negative real parts. In studies of parabolic partial differential equations, this is commonly referred to as the Crank-Nicholson method (see, e.g., ref. 12, p. 264). Another method which also has second-order accuracy for small $\lambda h$, but is not so well known, is given by the two step equation

$$v_{n+2} = \frac{1}{3} (4v_{n+1} - v_n + 2hv'_{n+2}) \tag{12}$$

Its use was suggested by Curtiss and Hirschfelder (ref. 1, eq. (17)), and it is basically the method used in boundary-layer-theory studies by Davis and Flügge-Lotz (see ref. 13). Equation (12) is not quite as accurate as equation (11), but it is more stable, that is, the parasitic eigenvalues are more heavily damped. Methods that are unconditionally stable for real negative values of $\lambda h$ and yet give higher order polynomial approximations for small

values of $\lambda h$ are given in both references 1 and 8. They can be derived by calculating the derivative $v'$ at $n + k$ using the values of $v$ at $n + k$, $n + k - 1$, . . ., $n$. The first five of these formulas are given in reference 14, pages 96-98.

The implicit method studied in this report is that given by equation (11). When applied to equation (4), there results the set of simultaneous equations

$$\left([I] - \frac{1}{2} h[A_n]\right)(\vec{w}_{n+1} - \vec{w}_n) = h\vec{F}_n + O(h^3) \tag{13}$$

where $[A_n]$ and $\vec{F}_n$ do not contain $s$ explicitly, and $\vec{F}_n = \vec{w}'_n$.

In order to compare this with other methods, it is necessary to estimate the time required to calculate the elements $\partial w'_i/\partial w_j$ of $[A_n]$ which, for nonlinear problems, must be reevaluated at each step. This time varies significantly with the problem and the details of the programming, but a reasonable general estimate for problems with $m$ simultaneous equations is to allow for $m$ additional time intervals equal to that required for evaluating $w'_n$. Roughly, one more such interval is required to solve the simultaneous equations, so the equivalent of about $m + 2$ calculations of $w'_n$ is required to advance the implicit method one step. Thus, approximately, $h = H(m + 2)/2$. The error of the implicit method is led by the term $(\lambda h)^3/12$ when applied to linear equations. For nonlinear equations the overall error is seen to remain at $O(h^3)$.

At this point the essential differences between the procedure proposed in this report and that proposed in reference 8 can be brought out. Basically, both processes depend upon an implicit method to insure stability. However, in reference 8 it is recommended that some of the equations (especially the fluid dynamic ones) be differenced by an explicit scheme. It was shown in the previous section that such a procedure invalidates the use of theorem I, which means that the stability and accuracy of the results become dependent on the size of the elements in the matrix. This can cause serious numerical difficulties. It is recommended here that <u>all</u> coupled equations be differenced at a given step by precisely the same differencing equation whether or not the method used is explicit or implicit. It is further suggested in reference 8 that higher order implicit methods with step numbers greater than 1 can be used. In view of the error introduced into equation (13) by the nonlinear terms, the employment of methods embedding polynomials of order greater than 2 cannot be justified in general. Further, the use of multistep methods detracts from the flexibility of the step-size adjustment available in one step methods. It is therefore proposed that when the use of an implicit method is warranted in the study of nonequilibrium flow problems, equation (13) is optimum.

## Locally Exact Methods

Finally, let us consider the approach to the problem used in reference 6. In methods such as this the basic equations are linearized to the form

given by equation (4), and these linearized equations are then solved by actually finding the eigenvalues of the matrix $[A_n]$. Except for the fact that these eigenvalues must be determined numerically, this provides exact solutions to the linear equations, and we will refer to such methods as "locally exact." The essential differences in the detailed use of the various methods is summarized in the following table. For a fixed accuracy, proceeding downward in the table corresponds to increased computing times.

| | |
|---|---|
| Conventional predictor-corrector methods | Must calculate $F_i$ in equation (2) for $i = 1, 2, 3, \ldots, m$ |
| Conventional Runge-Kutta[3] and Treanor methods | Must, in addition, calculate $\partial F_i/\partial u_i$ for $i = 1, 2, \ldots, m$ |
| Implicit methods | Must, in addition, calculate all off-diagonal terms $\partial F_i/\partial u_k$ for $i,k = 1, 2, \ldots, m$; $i \neq k$ |
| Locally exact methods | Must, in addition, calculate all eigenvalues of $[A_n]$ |

It is probably wise to estimate the largest local eigenvalue every once in a while in any event, but the exact calculation of all eigenvalues can be time-consuming, and, on this basis alone,[4] is to be avoided where possible. We seek, then, to find conditions under which the calculation of all the eigenvalues at every step can be avoided on the basis that such information increases neither the accuracy nor the stability of the results. In order to do this, consider the nature of our particular problem. Two possibilities occur. First, the linearized differential equations are themselves stable (inherent stability). That is, the real parts of all eigenvalues in $[A_n]$ are negative. Then, by theorem I, the implicit method given by equation (13) is <u>just as stable</u> as the locally exact methods, since the stability of both depends in the same way on exactly the same eigenvalues regardless of whether or not such eigenvalues are determined explicitly. Further, for general purposes, the implicit method is <u>just as accurate</u> as the locally exact methods since the order of error introduced by the linearization itself is as great as that caused by the differencing. Therefore, in the case of inherently stable differential equations, the continuous explicit calculation of the eigenvalues adds neither to the stability nor to the accuracy that is achieved by use of the implicit method given by equation (13).

The other possibility mentioned above comes about when the linearized equations are themselves unstable. It is conceivable in this case that the

---

[3]The standard, fourth-order, Runge-Kutta method is implied. For this and for Treanor's method $\partial F_i/\partial u_i$ can be determined by the calculations made at the intermediate step. In the Runge-Kutta method it is not necessarily used although sometimes it is employed to monitor step size.

[4]The experience of the authors has been that standard subroutines that provide eigenvalues for general real matrices with parasitic eigenvalues are <u>sometimes</u> (which makes it worse) of questionable accuracy with regard to the driving ones.

locally exact methods would remain stable for significantly larger step sizes than would be possible for the implicit one. This condition can occur if the initial conditions are just those which require the coefficients of the unstable terms in the locally exact solution to be zero. The use of locally exact methods to solve nonlinear problems with this kind of inherent instability would require justification in individual cases.

## Recommendations

On the basis of the above arguments, certain recommendations can be made which apply to the choice of a numerical method for integrating the equations (which are $m$ in number, including those due to the fluid mechanics) governing one-dimensional, nonequilibrium fluid flow. If $\lambda_1$ is the largest parasitic eigenvalue and $|\lambda_2|$ is the absolute value of the largest driving eigenvalue, and if, for accuracy, we require $|\lambda_2|h \leq \epsilon$, then, unless a special analysis of the particular problem indicates otherwise, it is strongly recommended that:

(a) All $m$ equations should be differenced by the same method for a given step,

(b) The method should be the implicit one given by equation (11) if $-\lambda_1/|\lambda_2| \gg (2/\epsilon)[(m + 2)/2]$

and two further, relatively weak, recommendations are:

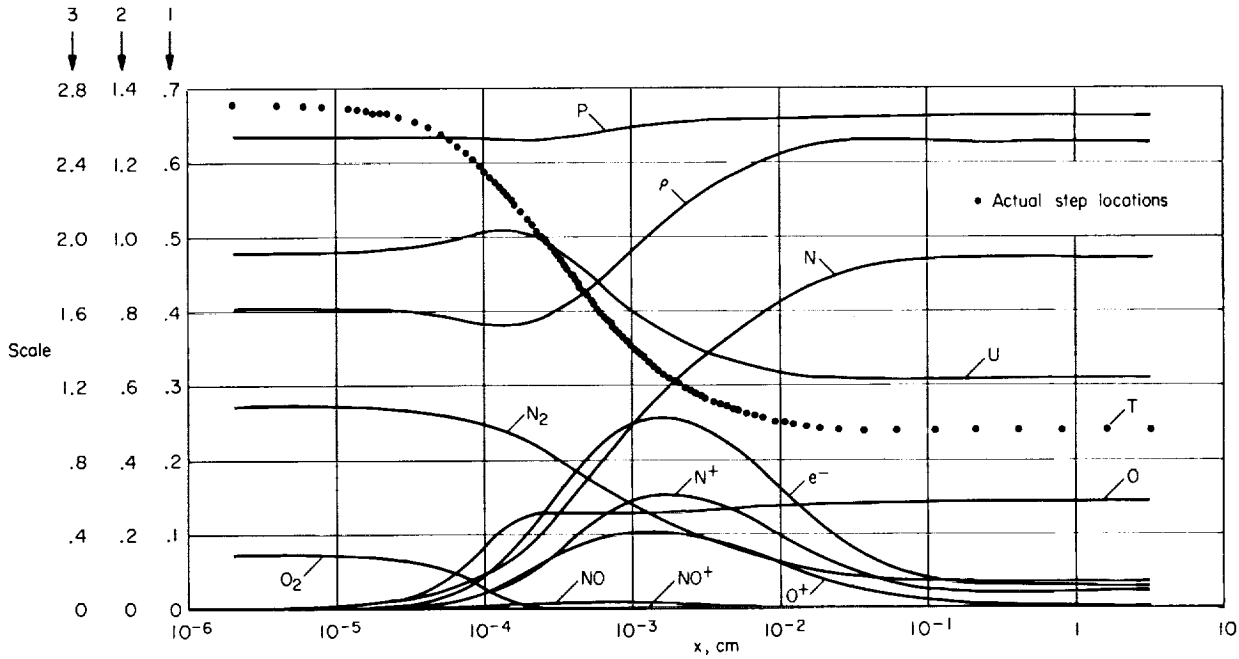
(c) The method can be the predictor-corrector combination given by equation (9) if $-\lambda_1/|\lambda_2| < (2/\epsilon)[(m + 2)/2]$

(d) The method can be the implicit one if $-\lambda_1/|\lambda_2| > (2/\epsilon)[(m + 2)/2]$

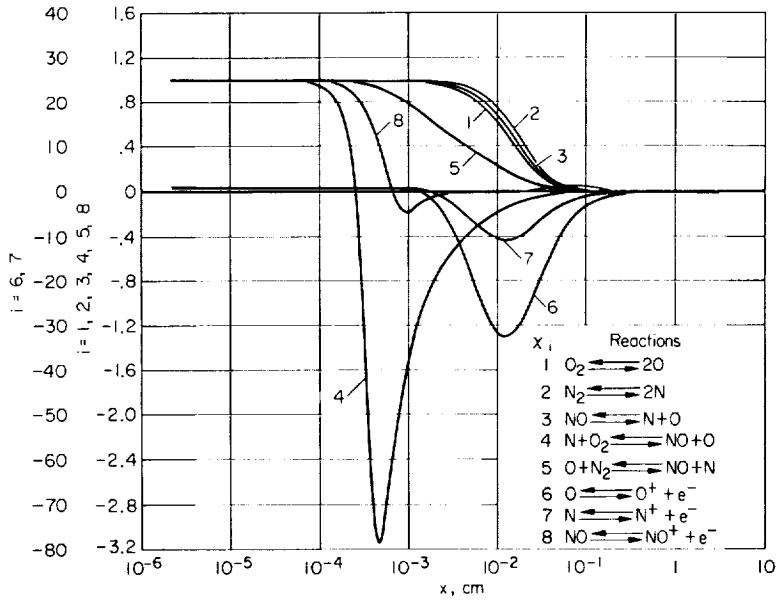
Clearly it is not the purpose of this report to extol the implicit method as a universal technique to be used in solving nonequilibrium problems. No numerical method is better than all other methods in all respects, each having its own merits. In modern machine languages most methods of the kind considered herein are quite easy to code and represent only a small part of a sophisticated program. In general, it appears that the option of using more than one of them, as the occasion demands, appears to be the really optimum procedure.

# COMPARISONS OF NUMERICAL SOLUTIONS FOR THE FLOW OF AIR IN CHEMICAL NONEQUILIBRIUM BEHIND A NORMAL SHOCK WAVE

All solutions were started by applying the Rankine-Hugoniot relations across a normal shock assuming a fixed chemical composition and local vibrational equilibrium for all species. The eight reactions shown in the insert in figure 2(b) were used. The rate equations and constants in them were the same as those used in reference 15 except that two reactions representing O and N ionization are included. The recombination rates of the latter were assumed to be 100 times those given in reference 16.




(a) Variation of physical variables.




(b) Variation of degree of nonequilibrium. Note change of scale for $X_8$ and $X_7$.

Figure 2.- Air flow behind a normal shock; $T_\infty = 300^\circ$ K, $p_\infty = 14{,}650$ dynes/cm$^2$, 9.144 km/sec. For scale see table I.

Equations (1) were analyzed for a variety of cases, a representative one of which is shown in complete detail. For the representative case the free-stream conditions are

$$\left.\begin{aligned} T_\infty &= 300^\circ \text{ K} \\ \rho_\infty &= 0.1695\times10^{-4} \text{ gm/cm}^3 \\ u_\infty &= 0.9144\times10^{6} \text{ cm/sec} \end{aligned}\right\} \tag{14}$$

The degree of nonequilibrium, $\chi_i$, for each reaction is defined as follows:

| | | |
|---|---|---|
| $\chi_1$ | $1 - \rho\gamma_1^2/\gamma_4 K_1$ | $O_2 \leftrightarrows 2O$ |
| $\chi_2$ | $1 - \rho\gamma_2^2/\gamma_5 K_2$ | $N_2 \leftrightarrows 2N$ |
| $\chi_3$ | $1 - \rho\gamma_1\gamma_2/\gamma_6 K_3$ | $NO \leftrightarrows N + O$ |
| $\chi_4$ | $1 - \gamma_1\gamma_6/\gamma_2\gamma_4 K_4$ | $N + O_2 \leftrightarrows NO + O$ |
| $\chi_5$ | $1 - \gamma_2\gamma_6/\gamma_1\gamma_5 K_5$ | $O + N_2 \leftrightarrows NO + N$ |
| $\chi_6$ | $1 - \rho\gamma_3\gamma_8/\gamma_1 K_6$ | $O \leftrightarrows O^+ + e^-$ |
| $\chi_7$ | $1 - \rho\gamma_3\gamma_9/\gamma_2 K_7$ | $N \leftrightarrows N^+ + e^-$ |
| $\chi_8$ | $1 - \gamma_3\gamma_7/\gamma_1\gamma_2 K_8$ | $N + O \leftrightarrows NO^+ + e^-$ |

where $K_i$ is the equilibrium constant for the ith reaction and $\chi_i$ varies from unity initially to zero as equilibrium is approached.

## Numerical Solutions Using the Implicit Method

The case described was computed over the range from just behind the shock to complete equilibrium, using the implicit method represented by equation (13). (The first few steps were explicit in order to establish the matrix elements.) The results are shown in figure 2, the scale for which is in table I. It should be remarked that the elements in $[A_n]$ were <u>not</u> found analytically. In every case they were calculated numerically by the equation

$$\frac{\partial F_i}{\partial w_j} \approx \frac{F_i(1.01w_j) - F_i(0.99w_j)}{0.02w_j} \tag{15}$$

which has an error $O[(0.02w_j)^2]$. It should also be remarked that considerable computing time can be saved if the paths in the subroutine which calculates $\partial F_i/\partial w_j$ are made different for u, $\rho$, T, and the species. As programmed, the

TABLE I.- VERTICAL SCALE PER MAJOR DIVISION IN FIGURES 2(a), 6, and 7

| Variable | Scale |
|---|---|
| u, cm/sec | $0.2\times10^{5}$ |
| $\rho$, gm/cm$^3$ | $.4\times10^{-4}$ |
| p, dynes/cm$^2$ | $.2\times10^{7}$ |
| T, $^{o}$K | $.4\times10^{4}$ |
| $O_2$, moles/gm | $.1\times10^{-1}$ |
| $N_2$, moles/gm | $.1\times10^{-1}$ |
| O, moles/gm | $.1\times10^{-1}$ |
| N, moles/gm | $.1\times10^{-1}$ |
| NO, moles/gm | $.1\times10^{-1}$ |
| $e^-$, moles/gm | $.1\times10^{-2}$ |
| $NO^+$, moles/gm | $.1\times10^{-2}$ |
| $O^+$, moles/gm | $.1\times10^{-2}$ |
| $N^+$, moles/gm | $.1\times10^{-2}$ |

machine computing time on an IBM 7094 averaged less than 90 seconds per case, and the total number of steps per case averaged less than 90.

While use of the implicit method over most of the range is far from optimum on the basis of computing time, the results permit us to test the reliability of the local expansion procedures used in equations (4) and (15) when applied for many steps over a highly nonlinear region. This test can be made since the first part of the solution (over half of the steps) can be repeated using the same step intervals and conventional predictor-corrector methods which require no local expansions or construction of $[A_n]$. Such calculations also permit us to evaluate the local eigenvalues over the entire flow region so that the numerical results can be interpreted and correlated with the analysis presented in the previous sections. Finally, such results permit us to see how the integration of a set of 13, nonlinear, coupled equations can proceed from a region having no parasitic eigenvalues into a region with several very large ones, without having stability difficulties.

Before comparing the results of the implicit method with those calculated by explicit means, let us discuss the technique used to govern the step size as the calculations proceeded.

## The Method Used to Control Step Size

Most techniques for adjusting the size of a step while the numerical integration is in progress depend upon the difference between a predicted and corrected value. One such method which tends to isolate the maximum negative eigenvalue in a set of stiff equations is presented in appendix B. When implicit methods are used, however, such a technique is not applicable since

predicted and corrected values are not calculated. Furthermore, since the implicit methods of interest are stable for all negative eigenvalues, the maximum eigenvalue is definitely not the one on which to base step size. As a result, the following method for controlling the step size of implicit methods was devised (it is not suitable for explicit methods when parasitic eigenvalues are present):

(1) After each step, compute $\Delta\vec{w} = |\vec{w}_{n+1} - \vec{w}_n|$.

(2) If any $\Delta w_j < 0.001$, ignore it in the following tests. (For the most part this limited the tests to the velocity and temperature variations.)

(3) If all $\Delta w_j$ that passed test 2 were such that $\Delta w_j/w_j < 0.01$, double the step size and take the next step. Otherwise, proceed to test 4.

(4) If any $\Delta w_j$ that passed test 2 was such that $\Delta w_j/w_j > 0.10$, halve the step size and take the next step. Otherwise, do not change the step size and proceed.

No other tests (e.g., negative species concentrations) were made. This method gave the step-size variation shown by the symbols on the temperature curves in figure 2(a).

## The Eigenvalue Distributions

A typical $[A_n]$ matrix (step no. 59) is shown in table II. The diagonal elements are underlined and the eigenvalues (obtained from a subroutine listed in ref. 17) are also given. Because of limitations in the subroutine, the eigenvalues below 17.2 are not to be trusted (calculated eigenvalues were never used in any of the numerical integrations). Such results were used to construct the curves shown in figure 3, which gives the largest positive eigenvalue and the largest negative eigenvalue throughout the region studied. The second and third largest negative eigenvalues over the latter part of the solutions are also included. The parasitic eigenvalues develop toward the end of the integration. At first glance, the region over which the parasitic eigenvalues exist seems relatively small. But, since the plots are logarithmic, it is actually relatively very large.

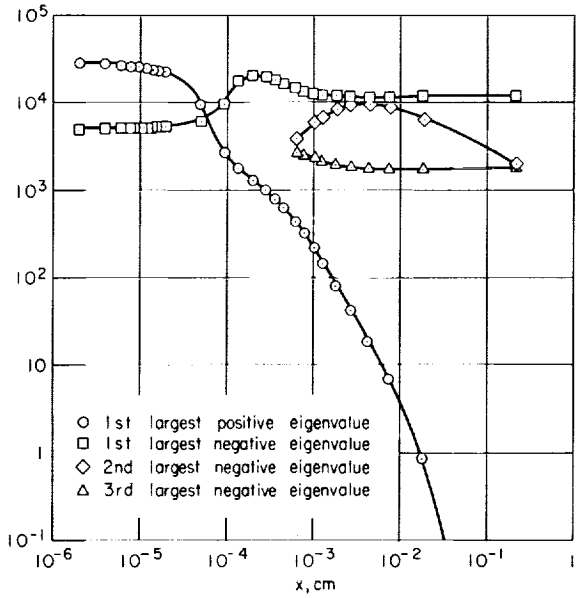



Figure 3.- Eigenvalues for solutions shown in figure 2.

TABLE II.- TYPICAL MATRIX $[A_n]$ AND EIGENVALUES

| Matrix | | | | | | | |
|---|---|---|---|---|---|---|---|
| Row \ Column | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 | -0.716E 02 | -0.481E 11 | -0.365E 04 | -0.234E 09 | -0.222E 09 | 0.180E 09 | 0.739E 09 |
| 2 | 0 | 0.206E 03 | 0.102E-04 | 0.655E 00 | 0.621E 00 | -0.504E 00 | -0.207E 01 |
| 3 | 0.300E 01 | -0.117E 11 | -0.100E 04 | -0.481E 08 | -0.563E 08 | 0.326E 08 | -0.118E 10 |
| 4 | -0.318E-06 | 0.187E 04 | -0.162E-04 | -0.613E 02 | 0.283E 02 | 0.543E 02 | 0.182E 05 |
| 5 | -0.780E-05 | 0.408E 05 | 0.266E-02 | 0.220E 03 | 0.226E 03 | 0.185E 03 | -0.590E 04 |
| 6 | 0.414E-07 | -0.908E 03 | 0.172E-03 | 0.627E 01 | 0.485E 01 | -0.836E 02 | -0.302E 01 |
| 7 | 0.802E-08 | -0.325E 02 | 0.214E-05 | 0.606E 01 | -0.284E 01 | -0.128E-00 | -0.121E 05 |
| 8 | 0.379E-05 | -0.197E 05 | -0.138E-02 | -0.134E 03 | -0.104E 03 | -0.604E 02 | -0.860E 02 |
| 9 | 0.236E-06 | -0.118E 04 | -0.460E-04 | 0.433E 02 | -0.237E 02 | -0.376E 01 | 0.607E 04 |
| 10 | 0.719E-08 | -0.374E 02 | 0.246E-04 | 0.486E 01 | 0.210E 01 | -0.269E 02 | -0.157E-00 |
| 11 | 0.583E-07 | -0.580E 03 | 0.333E-04 | 0.101E 01 | -0.111E 01 | -0.234E 02 | -0.487E 01 |
| 12 | -0.241E-07 | -0.290E 03 | 0.114E-03 | 0.410E-00 | 0.387E 01 | -0.333E 02 | 0.201E 01 |
| 13 | 0.986E-05 | 0.409E 04 | 0.755E-04 | 0.261E 02 | 0.293E 02 | 0.251E 02 | 0.344E 02 |
| Eigenvalues | | | | | | | |
| | -0.123E 05<br>0 | -0.782E 04<br>0 | -0.204E 04<br>0 | -0.948E 03<br>0 | -0.124E 03<br>0 | -0.806E 02<br>0 | -0.576E 02<br>0 |

| Matrix | | | | | | |
|---|---|---|---|---|---|---|
| Row \ Column | 8 | 9 | 10 | 11 | 12 | 13 |
| 1 | -0.524E 09 | -0.109E 10 | 0.134E 11 | 0.124E 09 | 0.119E 09 | -0 |
| 2 | 0.147E 01 | 0.305E 01 | -0.375E 02 | -0.347E-00 | -0.332E-00 | 0 |
| 3 | -0.125E 09 | -0.443E 09 | 0.206E 10 | 0.186E 08 | 0.194E 08 | -0 |
| 4 | -0.569E 02 | 0.165E 04 | 0.783E 04 | 0.738E 02 | 0.169E 02 | 0 |
| 5 | 0.532E 03 | 0.107E 04 | 0.798E 04 | 0.162E 03 | 0.204E 03 | 0 |
| 6 | -0.139E 01 | -0.243E 02 | -0.780E 04 | -0.571E 02 | -0.570E 02 | -0 |
| 7 | -0.388E-00 | 0.266E 03 | -0.422E-00 | -0.358E-00 | -0.343E-00 | -0 |
| 8 | -0.297E 03 | 0.528E 03 | -0.102E 03 | -0.735E 02 | -0.665E 02 | -0 |
| 9 | 0.599E 02 | -0.214E 04 | -0.173E 02 | -0.156E 02 | -0.151E 02 | -0 |
| 10 | -0.161E-00 | -0.153E-00 | -0.778E 04 | -0.134E-00 | -0.121E-00 | -0 |
| 11 | -0.210E 01 | -0.412E 02 | -0.415E 02 | -0.574E 02 | -0.100E 01 | -0 |
| 12 | 0.868E 00 | 0.170E 02 | 0.171E 02 | 0.458E-00 | -0.559E 02 | 0 |
| 13 | 0.351E 02 | 0.335E 02 | 0.411E 02 | 0.293E 02 | 0.264E 02 | 0 |
| Eigenvalues | | | | | | |
| | 0<br>0 | 0.101E 03<br>0 | 0.172E 02<br>0 | 0<br>0 | -0.381E-04<br>0 | 0<br>0 |

Figure 4 is a plot of $\lambda h$ in which $\lambda$ is the largest positive eigenvalue. The product $|\lambda h|_{max}$ in which $\lambda$ is the largest negative eigenvalue is shown in figure 5. These clearly establish the region outside of which numerical methods with limited stability boundaries would fail. For example, the second-order Runge-Kutta method (eq. (9)) would be unstable when $|\lambda h|_{max}$ exceeded 2.

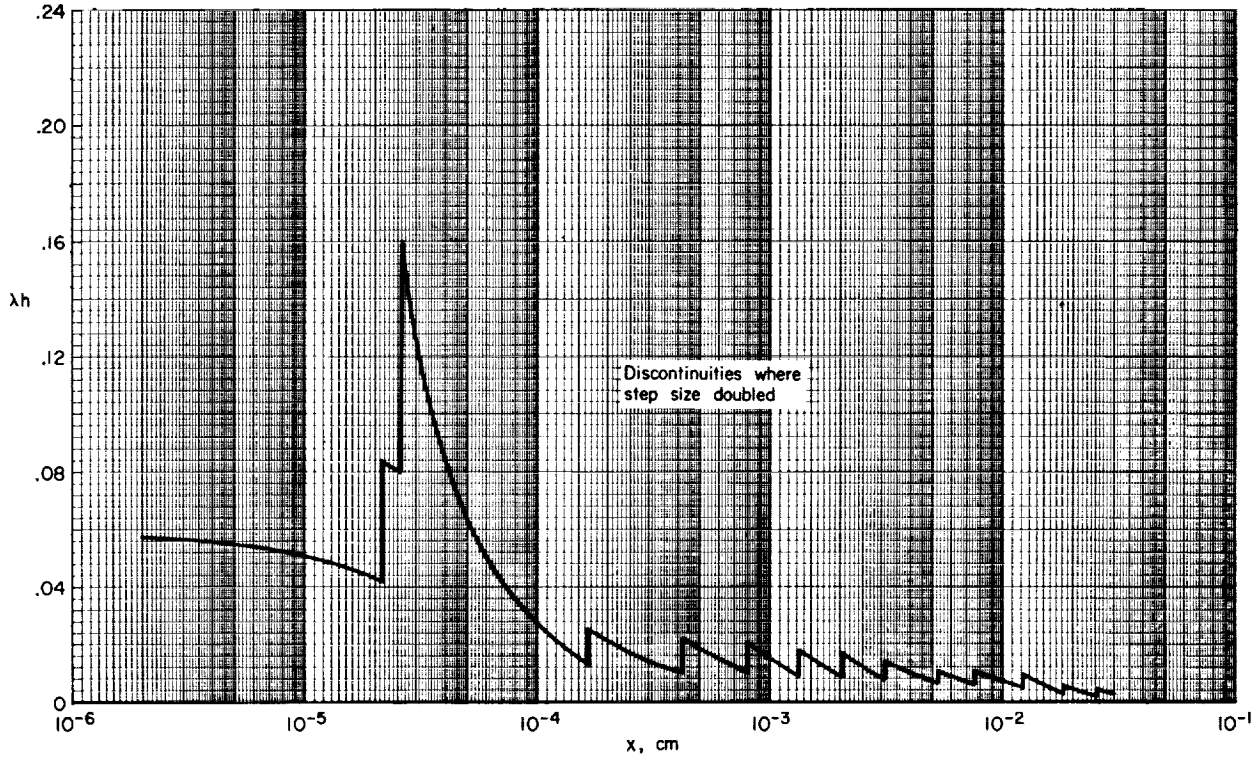



Figure 4.- Product of h and largest positive eigenvalue.

## Results Using Other Methods

The initial values given in equations (14) provided a model on which to base further numerical tests using different methods. First, the explicit second-order Runge-Kutta method given by equation (9) was combined with the method for step control given above and used to integrate the basic formulas directly in the form presented in equation (3). In the range where $|\lambda h|_{max}$ remained less than 2, the results were almost identical to those given by the implicit method applied to equation (4). This occurred for values of x less than about 0.002 (see fig. 5). A comparison of the two results is shown in table III. In each case, about 62 steps were taken and the variation of step size was essentially the same. (The slight difference in s is accounted for by the few, fixed, explicit steps initially used to develop a reliable $[A_n]$ matrix before the implicit method was turned on.) The results show that numerical methods that make use of

(a) Local expansion of nonlinear coupled equations,

(b) Numerical evaluation of the elements in $[A_n]$, and

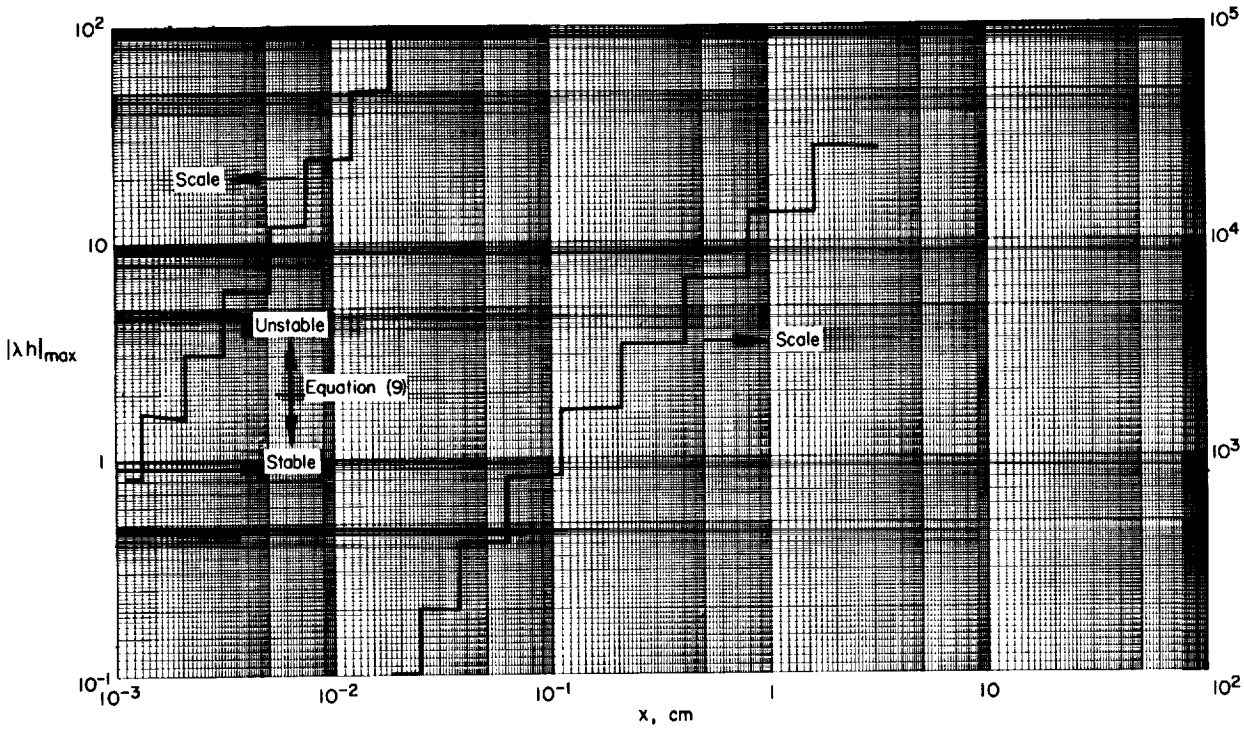



Figure 5.- Product of h and largest negative eigenvalue.

TABLE III.- COMPARISON BETWEEN RESULTS OF NUMERICAL INTEGRATIONS AFTER ABOUT 62 STEPS USING EXPLICIT AND IMPLICIT METHODS

| | Explicit (eq. (9)) | Implicit (eq. (11)) |
|---|---|---|
| s | 0.20140E-02 | 0.20100E-02 |
| x | .18588E-02 | .18560E-02 |
| u | .73188E 05 | .73179E 05 |
| ρ | .21177E-03 | .21181E-03 |
| T | .12326E 05 | .12327E 05 |
| O | .13125E-01 | .13122E-01 |
| N | .30181E-01 | .30177E-01 |
| $e^-$ | .25602E-02 | .25616E-02 |
| $O_2$ | .10907E-04 | .10903E-04 |
| $N_2$ | .11215E-01 | .11218E-01 |
| NO | .48791E-03 | .48814E-03 |
| $NO^+$ | .81259E-05 | .81074E-05 |
| $O^+$ | .10161E-02 | .10196E-02 |
| $N^+$ | .15359E-02 | .15339E-02 |

(c) Implicit integration of the equations thus derived

give an excellent numerical solution as far as accuracy is concerned, even when the spread of elements in $[A_n]$ is as drastic as that shown in table II. It should be remarked that all calculations were made in 8-place, floating-point arithmetic.

When the explicit method was used in the numerical integration for $x < 0.002$ and the implicit method for $x > 0.002$ (at all times monitoring the step size by the method for step control given above), the results at $x = 3.24$ were even closer to the "pure" implicit method than the comparison shown in table III. A plot of the complete results of these computations is shown in figure 6. The total machine time required for the combined explicit-implicit calculation was less than half of that required to calculate the entire range using the implicit method.

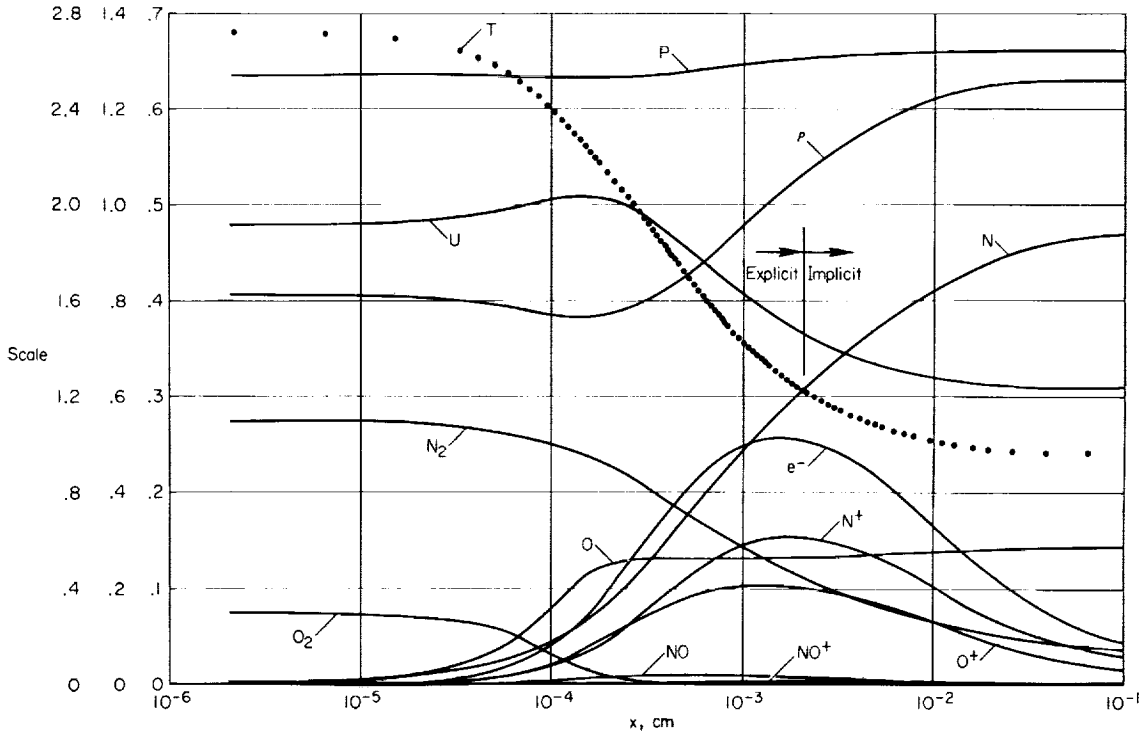
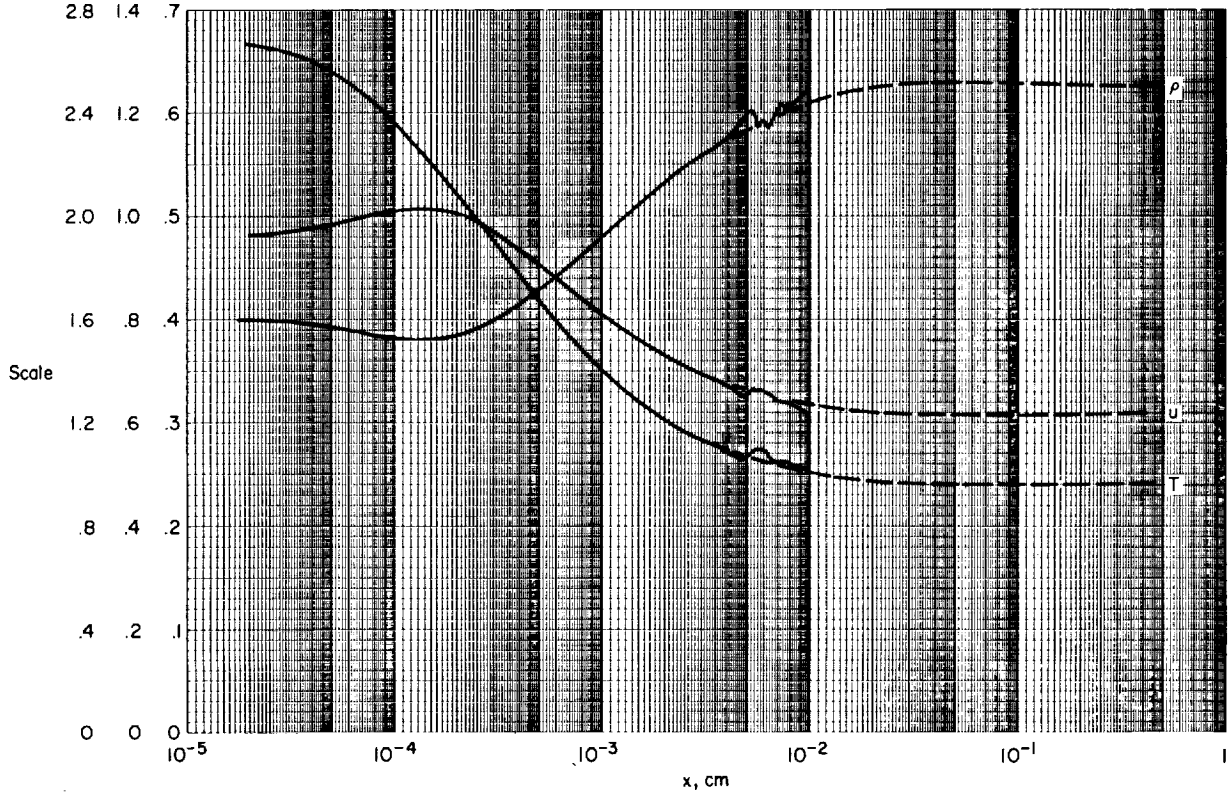



Figure 6.- Numerical calculation using the explicit method for $x < 0.002$ (62 steps) and implicit method for $x > 0.002$ (29 steps). Step control given on page 18.

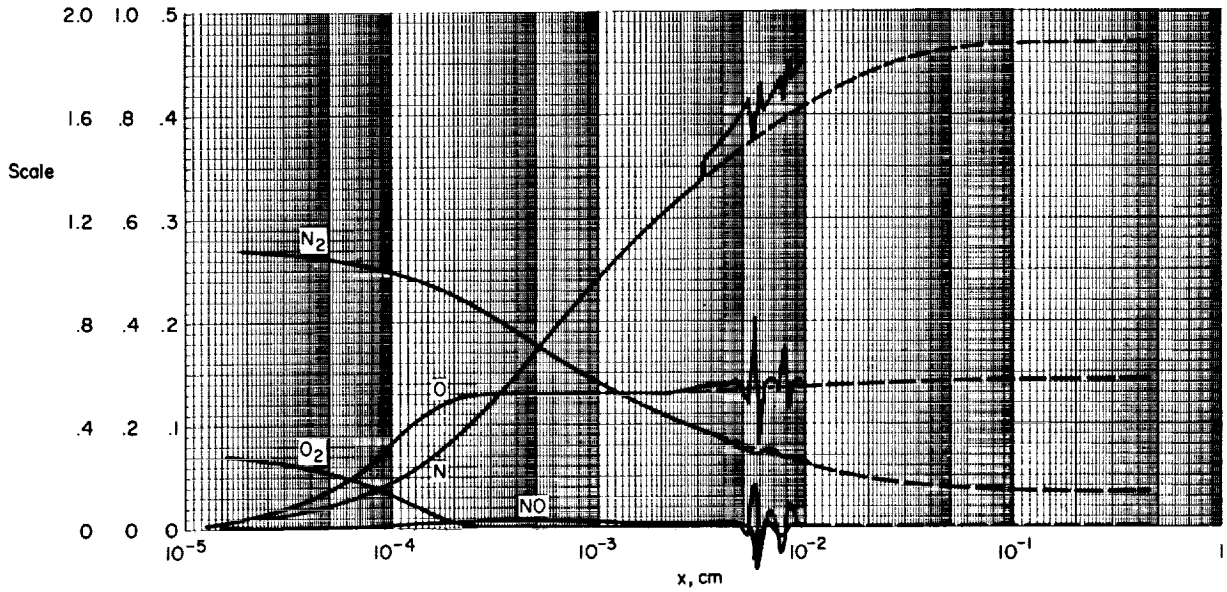
Another calculation was made using the explicit method given by equation (9) for $x < 0.002$, and then switching to Treanor's method for $x > 0.002$. The results are shown by the solid lines in figure 7. The method for step control given above was used throughout and $P_i$ was set to zero if it was calculated to be negative. The computation was continued for the same number of steps (29) as that used for the implicit continuation shown in figure 6. Under the imposed conditions, the results are what would be expected for the parasitic eigenvalue spread shown in figure 3 and the step control employed.

A final calculation, made using the computer program described in reference 15, is shown by the dashed curves in figure 7. For the lower values of $x$ these lines coincide with the solid lines. The computer program of reference 15 makes use of Treanor's method but with a suitably designed step control described in the reference. The program was made to compute for 5 minutes and progressed to a value of $x$ equal to 0.46, at which point the step size was 0.002. At every computed point the results agreed with those calculated by the implicit method and there were no visable oscillations. This shows that, with proper step control, Treanor's method can give excellent results. Its limitation in carrying solutions into regions with large parasitic eigenvalues is demonstrated by the fact that its running time was six times that required for the calculations shown in figure 6.

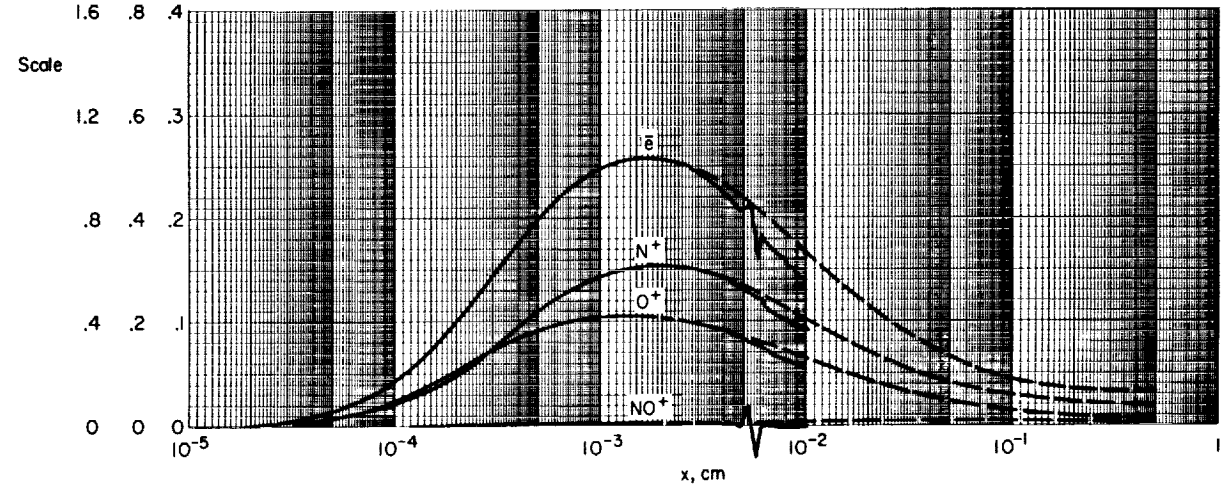



(a) Thermodynamic variables.

Figure 7.- Numerical calculation using equation (9) for $x < 0.002$ (62 steps) and Treanor's method - with the step control given on page 18 - for $x > 0.002$ (29 steps), solid lines. Numerical calculation using Treanor's method as programmed in reference 15, dashed lines.




(b) Molecules and atoms.




(c) Electrons and ions.

Figure 7.- Concluded.

## CONCLUDING REMARKS

If a set of ordinary differential equations can be linearized with regard to the dependent variables for a given value of the independent variable, then local behavior can be related to the eigenvalues of the matrix constructed from the linearized form. Then, if the differential equations are integrated numerically in a proper way, the stability (and accuracy, if the equations are autonomous) of the integration depends only upon these same eigenvalues and is independent of the detailed coupling of the set. When some of these eigenvalues are parasitic, conventional explicit methods can require excessive machine computing times. If properly applied, specially designed explicit methods, such as that due to Treanor, can give considerable reductions in computing time. However, we conclude from the results of this study, that if more than one very large parasitic eigenvalue occurs, the differential equations should be linearized locally and integrated implicitly (using the simplest implicit method compatible with the desired accuracy).

Ames Research Center
National Aeronautics and Space Administration
Moffett Field, Calif., 94035, May 31, 1967
129-04-03-02-00-21

## APPENDIX A

### EFFECT OF USING SEVERAL DIFFERENCING SCHEMES IN ONE STEP

In this appendix it is shown that when, for a given step, different schemes are used to difference the various differential equations in a coupled set, the stability and accuracy of the result will generally depend upon the individual elements in the [A] matrix of the differential set.

Consider the two linear differential equations

$$\left.\begin{aligned} w' &= a_{11}w + a_{12}v \\ v' &= a_{21}w + a_{22}v \end{aligned}\right\} \tag{A1}$$

Apply the Euler predictor

$$u_{n+1} = u_n + hu_n' \tag{A2}$$

to the first, and the implicit modified Euler method

$$u_{n+1} = u_n + \frac{1}{2}h(u_{n+1}' + u_n') \tag{A3}$$

to the second. There results

$$w_{n+1} = w_n + h(a_{11}w_n + a_{12}v_n)$$

$$v_{n+1} = v_n + \frac{1}{2}h(a_{21}w_{n+1} + a_{22}v_{n+1} + a_{21}w_n + a_{22}v_n)$$

Introduce the operator $E \equiv \exp[h(d/dx)]$ and combine terms. We find the matrix equation

$$\begin{bmatrix} E - (1 + a_{11}h) & -a_{12}h \\ -\frac{1}{2}a_{21}h(E + 1) & E - 1 - \frac{h}{2}a_{22}(E + 1) \end{bmatrix} \begin{bmatrix} w_n \\ v_n \end{bmatrix} = 0$$

having a characteristic equation which can be written

$$\det \begin{pmatrix} a_{11} - S_1(E,h) & a_{12} \\ a_{21} & a_{22} - S_2(E,h) \end{pmatrix} = 0 \tag{A4}$$

in which $S_1(E,h)$ and $S_2(E,h)$ are the operators

$$S_1 = \frac{E - 1}{h}, \qquad S_2 = \frac{2(E - 1)}{h(E + 1)} \tag{A5}$$

It is at once evident that the form of $S_1$ results entirely from the choice of an Euler predictor for the first row, and the form of $S_2$ is brought about entirely by the use of an implicit modified Euler method for the second row. The generalization is clear. If a variety of methods are used on the individual equations, a variety of operators appear on the diagonal of [A].

Returning to the simple example given by equation (A4), and using the identities

$$\lambda_1 + \lambda_2 = a_{11} + a_{12}$$

$$\lambda_1\lambda_2 = a_{11}a_{22} - a_{12}a_{21}$$

where $\lambda_1$ and $\lambda_2$ are the eigenvalues of

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

we can reduce equation (A4) to

$$S_1S_2 - (\lambda_1 + \lambda_2)S_2 + a_{22}(S_2 - S_1) + \lambda_1\lambda_2 = 0 \tag{A6}$$

Suppose the differencing equations (A2) and (A3) had been the same. Then $S_1 = S_2 = S$ and the left side of equation (A6) factors into $(S - \lambda_1)(S - \lambda_2)$ which shows no dependence on any of the elements, $a_{ij}$, regardless of the functional dependence of $S$ on $E$ and $h$. Obviously this would be true in general. However, for the different values of $S_1$ and $S_2$ in the example, equation (A6) becomes

$$E^2\left(1 - \frac{ha_{22}}{2}\right) + E\left[-2 - h(\lambda_1 + \lambda_2) + \frac{h^2}{2}\lambda_1\lambda_2 + ha_{22}\right]$$

$$+ \left[1 + h(\lambda_1 + \lambda_2) + \frac{h^2}{2}\lambda_1\lambda_2 - \frac{ha_{22}}{2}\right] = 0 \tag{A7}$$

which clearly depends on the element $a_{22}$ as well as on the eigenvalues.

As a simple illustration, let $\lambda_1 = 1$, $\lambda_2 = -1$, and $\tau = (1/2)ha_{22}$. The roots to equation (A7) are then

$$E = 1 \pm h\frac{\sqrt{1 - \frac{1}{2}\tau}}{1 - \tau} + O(h^2)$$

The two matrices

$$\begin{bmatrix} 17 & -20 \\ 14.4 & -17 \end{bmatrix}, \quad \begin{bmatrix} -17 & 14.4 \\ -20 & 17 \end{bmatrix}$$

have the same eigenvalues (+1 and -1), but have values of $\tau$ equal in one case to $-8.5h$ and to $+8.5h$ in the other. Use of the mixed differencing schemes in the example is clearly inadvisable.

## APPENDIX B

## ANALYSIS OF THE PARAMETER P IN TREANOR'S METHOD

The ratio, $P_i$, between the difference in the predicted and corrected values of the derivative and the function calculated at the intermediate step in a standard, fourth-order, Runge-Kutta method is a basic parameter used in the method proposed by Treanor. In this appendix we analyze the significance of $P_i$ when it is calculated from linear equations with constant coefficients. A set of "representative" stiff equations is then introduced and analyzed by a variety of methods. It is shown that for any method an effective $P_i$ (that is, a number equal to that which would have resulted if the Runge-Kutta method had been used) can be calculated as the integration proceeds. The consequence of the results with regard to the accuracy and possible modification of Treanor's method is discussed and a variety of examples are shown.

The predictor-corrector equivalent of the standard, fourth-order, Runge-Kutta method is

$$\left.\begin{aligned} v^{(1)}_{n+(1/2)} &= v_n + \left(\frac{h}{2}\right) v_n' \\ v^{(2)}_{n+(1/2)} &= v_n + \left(\frac{h}{2}\right) v^{(1)'}_{n+(1/2)} \\ v^{(3)}_{n+1} &= v_n + 2\left(\frac{h}{2}\right) v^{(2)'}_{n+(1/2)} \\ v_{n+1} &= v_n + \frac{1}{3}\left(\frac{h}{2}\right)\left[v^{(3)'}_{n+1} + 2\left(v^{(2)'}_{n+(1/2)} + v^{(1)'}_{n+(1/2)}\right) + v_n'\right] \end{aligned}\right\} \quad \text{(B1)}$$

where the superscript designates a family generated in a single step. Let $v$ represent the $i$th element in the vector $\vec{w}$. In Treanor's method the value of $P_i$ used in equation (10) is determined by the equation

$$P_i = -\frac{v^{(2)'}_{n+(1/2)} - v^{(1)'}_{n+(1/2)}}{v^{(2)}_{n+(1/2)} - v^{(1)}_{n+(1/2)}} \quad \text{(B2)}$$

Consider the coupled set of linear differential equations with constant coefficients

$$\vec{w}' = \frac{d\vec{w}}{ds} = [A]\vec{w}(s) + \vec{f}(s) \quad \text{(B3)}$$

which are not autonomous, and in which [A] is a matrix of constants. The explicit results of applying equations (B1) to (B3) are, for the first family,

$$\left.\begin{aligned} \vec{w}^{(1)}_{n+(1/2)} &= \left([I] + \left(\frac{h}{2}\right)[A]\right)\vec{w}_n + \left(\frac{h}{2}\right)\vec{f}_n \\ \vec{w}^{(1)\prime}_{n+(1/2)} &= \left([A] + \left(\frac{h}{2}\right)^2 [A]^2\right)\vec{w}_n + \left(\frac{h}{2}\right)[A]\vec{f}_n + \vec{f}_{n+1} \end{aligned}\right\} \tag{B4}$$

and, for the second family,

$$\left.\begin{aligned} \vec{w}^{(2)}_{n+(1/2)} &= \left([I] + \left(\frac{h}{2}\right)[A] + \left(\frac{h}{2}\right)^2 [A]^2\right)\vec{w}_n + \left(\frac{h}{2}\right)^2 [A]\vec{f}_n + \left(\frac{h}{2}\right)\vec{f}_{n+1} \\ w^{(2)\prime}_{n+(1/2)} &= \left([A] + \left(\frac{h}{2}\right)[A]^2 + \left(\frac{h}{2}\right)^2 [A]^3\right)\vec{w}_n + \left(\frac{h}{2}\right)^2 [A]^2\vec{f}_n + \left(\frac{h}{2}\right)[A]\vec{f}_{n+1} + \vec{f}_{n+1} \end{aligned}\right\} \tag{B5}$$

From these we form the ratio

$$\vec{P}_n = -\frac{[A]\left\{[A]^2\vec{w}_n + [A]\vec{f}_n + \dfrac{\vec{f}_{n+(1/2)} - \vec{f}_n}{h/2}\right\}}{\left\{[A]^2\vec{w}_n + [A]\vec{f}_n + \dfrac{\vec{f}_{n+(1/2)} - \vec{f}_n}{h/2}\right\}} \tag{B6}$$

where the division is defined to mean that an element in the vector in the numerator is divided by the corresponding element of the vector in the denominator.

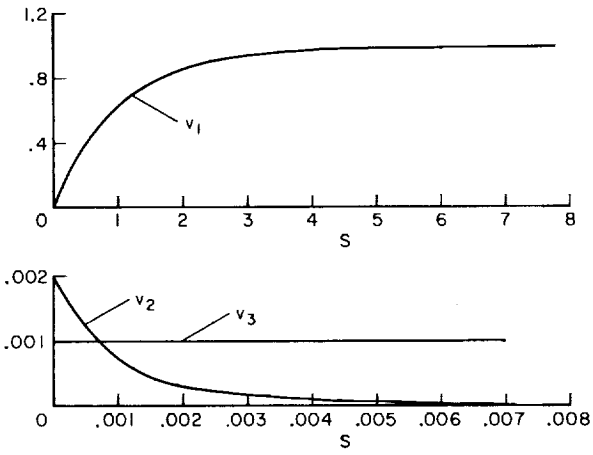
In studying the effect of parasitic eigenvalues, as we have defined them, we can consider the dominant eigenvalues of [A] to be real and distinct. Now let $\vec{g}$ be any linear combination of the eigenvectors of [A]. A well-known result under these conditions (see, e.g., ref. 9, pp. 205 and 206) is that if

$$\vec{\Lambda} = \frac{[A]^j\vec{g}}{[A]^{j-1}\vec{g}} \tag{B7}$$

and division is defined as above, when $j$ is increased, all elements in $\vec{\Lambda}$ approach the dominant eigenvalue whose eigenvector is given a nonzero weight in the formulation of $\vec{g}$.

We now seek to find how closely the results just presented apply to the value of $\vec{P}_n$ as $n$ is advanced in a numerical solution. Our procedure is to make a simple numerical experiment starting with three uncoupled equations and their initial values such that

$$\left.\begin{aligned} v_1' &= -v_1 + 1 \quad , \quad v_1(0) = 0 \\ v_2' &= -500v_2 \quad , \quad v_2(0) = 0.002 \\ v_3' &= -1000v_3 + 1 \quad , \quad v_3(0) = 0.001 \end{aligned}\right\} \tag{B8}$$




Sketch (a).- Exact solutions of equations (B8).

The exact solutions are shown in sketch (a), and the value of [L] as used in equation (5) is given by

$$[L] = [I] \begin{bmatrix} -1 \\ -500 \\ -1000 \end{bmatrix}$$

The eigenvalue -1000 is everywhere parasitic since $v_3$ does not vary at all, and the eigenvalue -500 quickly becomes parasitic relative to -1. These equations are coupled by a set of transformations given in equations (6) and (7); the matrix used for [B] (chosen at random) is

$$[B] = \begin{bmatrix} -1.8944 & -0.84661 & -1.2519 \\ -.14343 & -.48809 & -.40617 \\ 1.4721 & -1.1597 & -.41521 \end{bmatrix} \tag{B9}$$

The corresponding matrix [A], which equals $[B][L][B]^{-1}$, is[1]

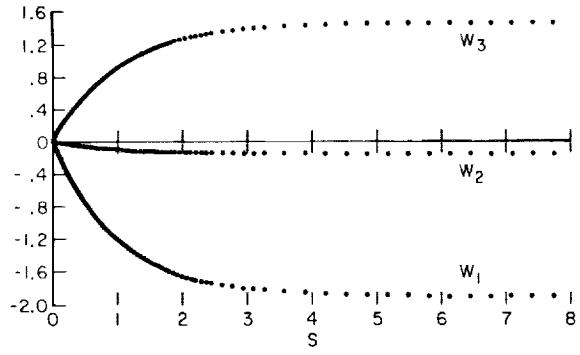
$$[A] = \begin{bmatrix} -19409 & 74820 & -17686 \\ -4657.6 & 17717 & -4267.3 \\ 314.68 & -2187.2 & 190.85 \end{bmatrix} \tag{B10}$$

The coupled equations for $\vec{w}$ take the form

---

[1]All computations were made in double precision to isolate roundoff errors. We present [A] only to illustrate the spread in the values of the elements, and both [A] and [B] have been truncated, for presentation, to five places.

$$\vec{w}' = [A]\vec{w} + [B]\vec{f} \qquad \text{(B11)}$$




Sketch (b).- Variation of $\vec{w}$ in equation (B11) calculated by the implicit method using step control (B12).

The behavior of the exact solutions for $\vec{w}$ is shown in sketch (b). Several numerical methods were used to integrate equation (B11) for about 100 steps, and for each method the variation of $\vec{P}$ was recorded.

Before discussing the results, it is necessary to say something about the step-size control used in the integration. In cases with parasitic eigenvalues, the magnitude of the largest negative eigenvalue is not necessarily the best parameter with which to judge step size. A variety of methods for controlling step size has been proposed, but we do not wish to list or judge them here. The particular procedure adopted in all the numerical methods discussed in this appendix is the following:

(1) Choose an initial step size.

(2) Advance one step and compute

$$\Delta w = \max\left|\vec{w}_{n+1} - \vec{w}_n\right|$$

(3) If $\Delta w > 0.05$ halve the step size and repeat.

(4) If $\Delta w < 0.0125$ double the step size and repeat unless

(a) An imposed maximum has been reached,

(b) The step size was previously doubled in this same step.

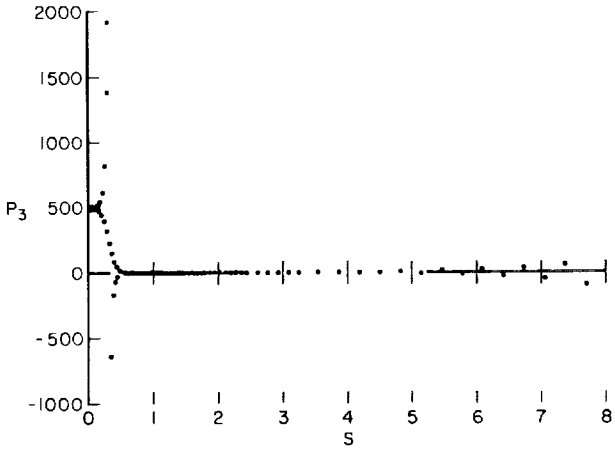
(B12)

The study of equation (B11) is a controlled experiment and this procedure was designed especially for it. The procedure is <u>not</u> recommended in general.

Using the implicit method given by equation (11) together with the step-size control just described, the equations for $\vec{w}$ were integrated for 100 steps starting with a step size equal to 0.02. The results, plotted in sketch (b), were excellent for the full 100 steps. The value of s was advanced to 7.74 (the symbols show the actual step locations) and the final 14 step sizes were equal to 0.32, the maximum allowed. When, after 100 steps, $\vec{w}$ was uncoupled (using the relation $\vec{v} = [B]^{-1}\vec{w}$), it compared to the exact analytical solution for $\vec{v}$ as follows:

| | Exact | Numerical |
|---|---|---|
| $v_1$ | 0.999565 | 0.999582 |
| $v_2$ | .000000 | .000000 |
| $v_3$ | .001000 | .001000 |

Under the constraints imposed by the method used for step control, the step size remained at 0.02 for the first 55 steps. During the first 10 of these, $\vec{P}$ was essentially the same for all three w's and equal to about 500. A transition occurred between steps 10 and 40 during which wild fluctuations[2] of $\vec{P}$ were noted. From step 40 through step 73 all values of $\vec{P}$ were again very uniform and all remained equal to about 1. At this point the step size was doubled from 0.04 to 0.08 and, as further doubling occurred and the w's approached their asymptotic values, $\vec{P}$ began to fluctuate. The actual variation of $\vec{P}$ is shown in sketch (c).



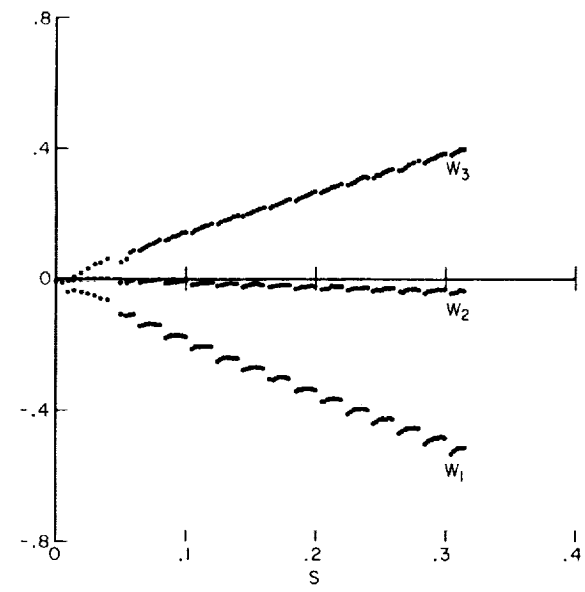

Sketch (c).- Value of $P_3$ calculated from equation (B6) using results from implicit method. Variations of $P_1$ and $P_2$ were nearly identical with $P_3$.

With the result given by (B6) and the discussion concerning (B7), the variation of $\vec{P}$ shown in sketch (c) can be explained. In the first place we notice that the dominant eigenvalue (the one associated with $v_3$ and equal to -1000) was never detected by equation (B6). The reason is simple. The exact solution for $v_3$ under the given initial conditions is a constant. In the terminology used to discuss equation (B7), the corresponding eigenvector has a zero weight in the construction of $\vec{w}$. To be sure truncation errors will begin to excite $v_3$ but, since the method is absolutely stable for all negative $\lambda$, this excitation is continually suppressed and has no effect on the value of $\vec{P}$ throughout most of the calculation. On the other hand, the eigenvector associated with $v_2$ is activated by the initial conditions and, for a while, appears in all three elements of $\vec{P}$. But relative to $v_1$, $v_2$ is very heavily damped and in any numerical integration that is stable for all negative $\lambda$, its influence eventually disappears from the leading significant figures that affect $\vec{P}$. For the implicit method used, this disappearance is completed when $s \approx 0.6$ and, at that point, the largest eigenvalue detectable in $\vec{w}$ is -1. The disappearance of $v_2$ does not occur at once, though, and for many steps a "struggle for dominance" (ref. 9, p. 206) between the eigenvalues associated with $v_1$ and $v_2$ takes place. This phenomenon is quite evident in sketch (c) for values of s around 0.3. Eventually all elements of $\vec{w}$ approach a constant and, although the solution itself is quite stable, $\vec{P}$ begins a meaningless fluctuation brought about by small truncation errors in all of the terms.

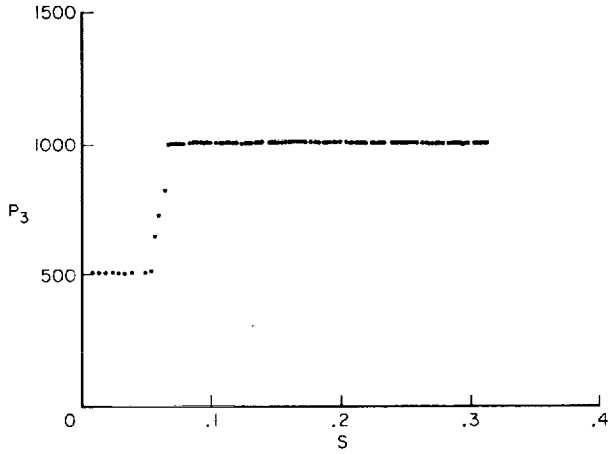
Next, equation (B11) was analyzed by a standard, fourth-order Runge-Kutta method using the identical step-size control. The results for $\vec{w}$ are shown in sketch (d). The method is unstable for $|\lambda h| > 2.8 (|\lambda H| > 1.4)$ so,

---

[2]It should be emphasized that $\vec{P}$ was in no way used in the numerical integration.




Sketch (d).- Solution of equations (B11) using standard, fourth-order, Runge-Kutta method with step-size control (B12).




Sketch (e).- Value of $P_3$ calculated from equation (B6) or equation (B2) using results of Runge-Kutta method. Variations of $P_1$ and $P_2$ are nearly identical with $P_3$.

with this kind of step control, it automatically settled[3] on a step size equal to 0.0025 for most of the integration. The scallop effect (which can also be noticed in the Runge-Kutta results presented in ref. 3) occurs because every once in a while, for just one step, use of the unstable step size, 0.005, is permitted by the particular test employed. This causes a relatively large error in $\vec{w}$, due to the presence of $v_3$. Immediately the method is forced to return to its stable step size and remain there until this error is sufficiently reduced, at which point the phenomenon is repeated. The value of s is advanced to 0.315 in 100 steps, and for this s the uncoupled values[4] of $\vec{v}$ are

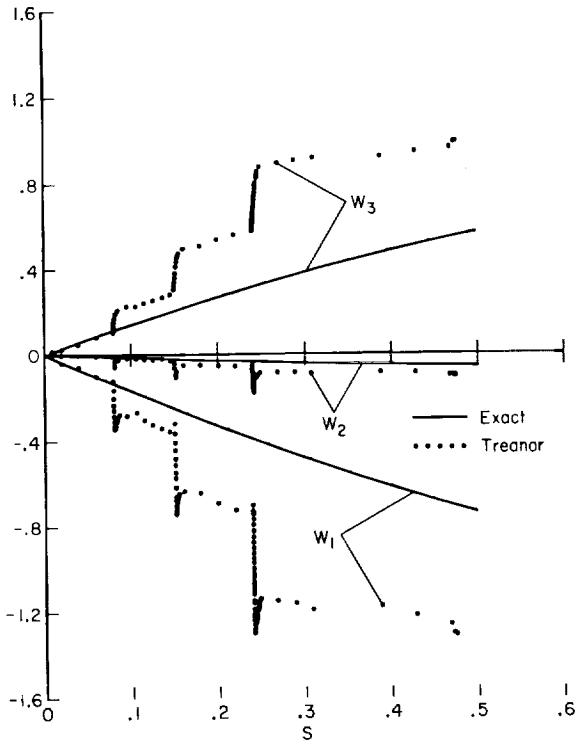
| | Exact | Numerical |
|---|---|---|
| $v_1$ | 0.270211 | 0.270211 |
| $v_2$ | .000000 | .000000 |
| $v_3$ | .001000 | .007086 |

The value of $\vec{P}$ generated by this method (once again, this value was never used in the integration) is shown in sketch (e). The method started itself with step sizes varying between 0.02 and 0.005 because it could not immediately detect the largest negative eigenvalue. After 13 steps, however, the error introduced by the unstable integration brought the presence of $v_3$ clearly into view, and throughout the rest of the calculation all elements of $\vec{P}$ were about 1000.

Next, Treanor's method with the same step-size control was used to integrate equation (B11). This time, of course, $\vec{P}$ was both calculated and used at each step. The results for $\vec{w}$ are shown in sketch (f). After the third step the method had "found" the largest eigenvalue and all three

[3]The step-size control being used is quite inefficient for this and other methods with limited stability boundaries. In the presence of parasitic eigenvalues such methods increase their step size until they reach their stability boundaries and then proceed to first double and then halve at almost every step.

[4]If the uncoupling is made just after an unstable step size is used, the error in $v_3$ is about five times that shown.




Sketch (f).- Solution of equations (B11) using Treanor's method with step-size control (B12).

elements of $\vec{P}$ were nearly 1000 for steps 3 through 5. In this range a step size of 0.02 passed all tests. This procedure is unstable for the -500 eigenvalue, however, and at the sixth step $P_1$ suddenly became negative, while $P_2$ and $P_3$ changed to 97 and 364, respectively. The rule in using Treanor's method is to set negative values of $P_j$ equal to zero and proceed. This can cause an abrupt change in the accuracy and stability of the integration, a change which, in general, depends critically (see appendix A) on the magnitude and sign of the individual elements in [A]. In the present example the result was the same as if a completely new set of initial values were introduced, and the integration departed drastically from the desired solution. At each of the three abrupt vertical rises in $w_3$ shown in sketch (f), the sign of one of the elements in $\vec{P}$ suddenly became negative while the other two remained at large positive values. It was further noted that a few steps before each of the less abrupt, nearly horizontal turns the same behavior in $\vec{P}$ occurred.

To give Treanor's method a chance, it was again used to integrate equation (B11), but with a step size fixed at 0.008, and <u>all elements of</u> $\vec{P}$ <u>fixed</u> at 1000. According to the stability boundaries shown in figure 1, this integration should be stable throughout. After 100 steps s had been advanced to 0.8 and the uncoupled values of $\vec{v}$ were

| | Exact | Numerical |
|---|---|---|
| $v_1$ | 0.550671 | 0.550576 |
| $v_2$ | .000000 | .000001 |
| $v_3$ | .001000 | .001000 |

To verify further the stability boundaries in figure 1, this type of run was repeated with all elements of $\vec{P}$ again fixed at 1000, but with a fixed step size equal to 0.01. The results were an interesting verification of theorem I and corollary 2. After 50 steps s was advanced to 0.5 and when $\vec{w}$ was uncoupled, there resulted for $\vec{v}$

| | Exact | Numerical |
|---|---|---|
| $v_1$ | 0.399504 | 0.399547 |
| $v_2$ | .000000 | $.377562 \times 10^{9}$ |
| $v_3$ | .001000 | .001013 |

All elements of the computed[5] $\vec{P}$ (which, of course, was not used) were almost exactly 500 throughout the run. Notice that, although an enormous error is made in $v_2$ (all elements of $\vec{w}$ were greater in magnitude than $10^8$ at the fiftieth step), the two uncoupled terms for which the method is stable remained accurate through the fourth significant figure.

The results of the test runs made using Treanor's method with fixed values of $P_j$ and step size suggested the following modifications to his method which would make it more reliable in general cases. One modification would be to

(1) Use the same differencing scheme

(2) Find $\vec{P}$ by means of equation (B2) and let $P_{max} = \max(\vec{P})$

(3) If $P_{max}$ is negative, set $P_{max} = 0$

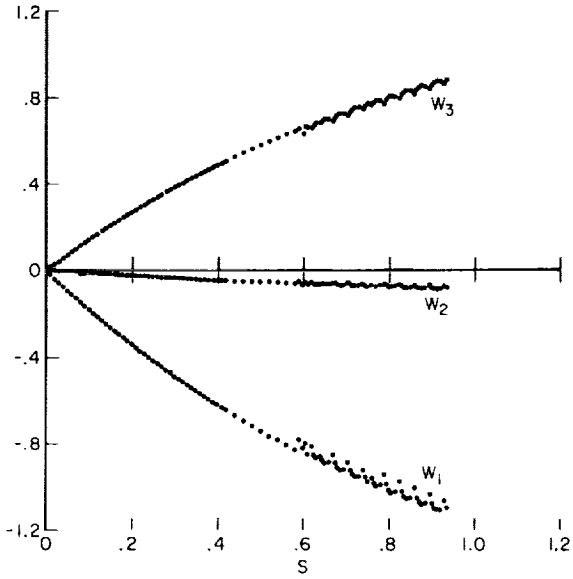
(4) Reset all elements of $\vec{P}$ with $P_{max}$ so that all equations are differenced by the same equations in a given step.

When applied to equations with widely different parasitic eigenvalues, this method has a real upper stability boundary $|\lambda h| = 10$ and $|\lambda H| = 5$.

This modified method was combined with the same step control employed in the previous examples and used to integrate equation (B11). The results for $\vec{w}$ are shown in sketch (g). Since it was limited by the unrefined doubling-halving step control procedure, the method was forced to fluctuate between step sizes of 0.005 and 0.01 for which it was stable and unstable, respectively. The scalloped effect caused when this kind of step control is used in methods with bounded stability is evident. The computed values of $\vec{P}$ are shown in sketch (h). In each step all three values of P were made to equal the maximum one shown. The method advanced s to 0.935 in 100 steps. Define the error terms

$$\Delta v_j = (v_j)_{exact} - (v_j)_{numerical}$$

and $\Delta v_2$ and $\Delta v_3$ ($\Delta v_1 < 0.0002$ throughout) are plotted in sketch (i).




Sketch (g).- Treanor's modified method (B13) with step-size control (B12) applied to equations (B11).

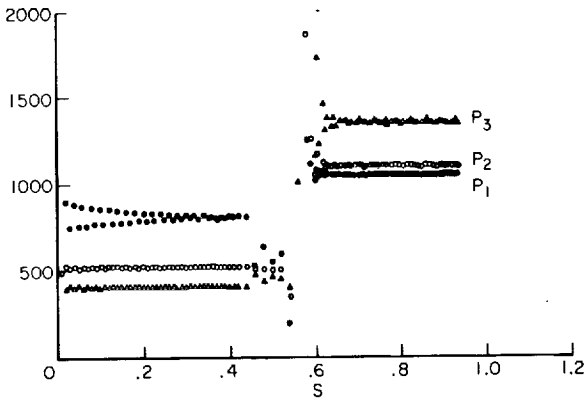
A second modification to Treanor's method would be to

(1) Use the same differencing scheme

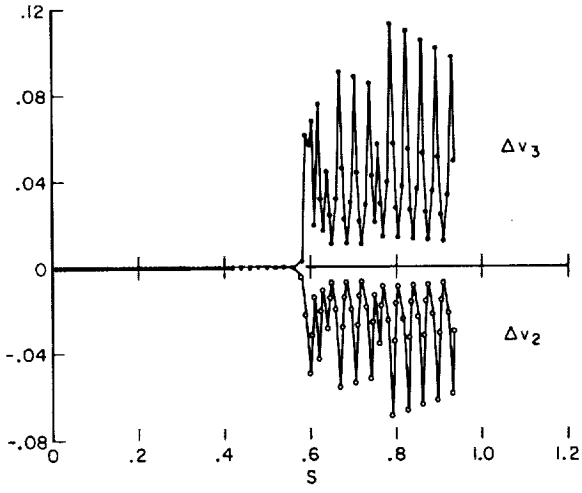
[5]Compute either by equation (B2) or (B6), the results being the same.

(2) Set all elements of $\vec{P}$ equal to $8/h$ at every step.

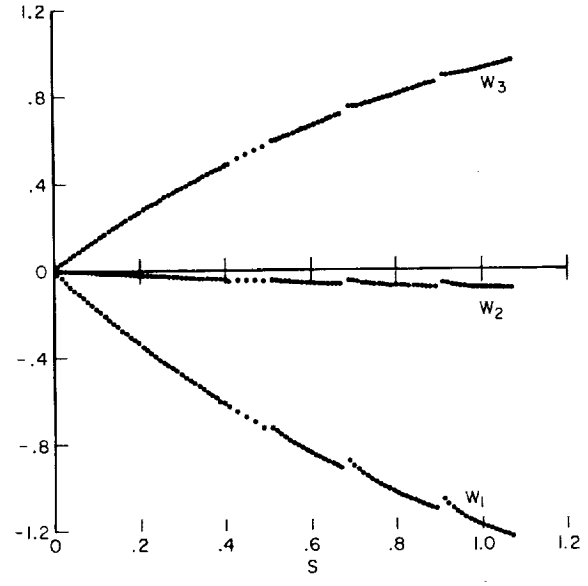
Such a method also has a real stability boundary $|\lambda H| = 0.5$. The results of applying it to equation (B11) are shown in sketches (j), (k), and (l). When



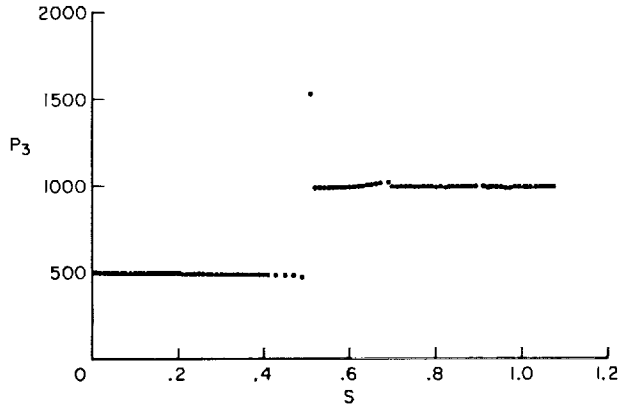

Sketch (h).- Values of $\vec{P}$ calculated using Treanor's modified method (B13).



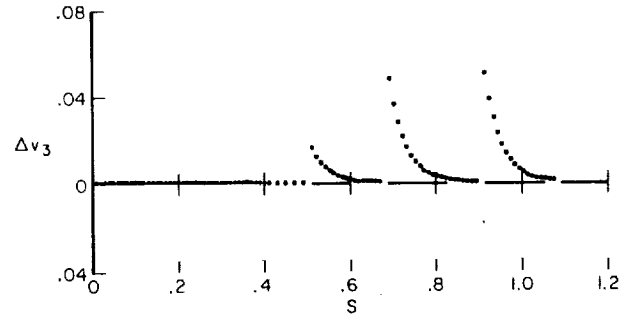

Sketch (i).- Errors in $v_2$ and $v_3$ using Treanor's modified method (B13).




Sketch (j).- Treanor's modified method (B14) with step-size control (B12) applied to equations (B11).




Sketch (k).- Value of $P_3$ calculated using Treanor's modified method (B14). Variations of $P_1$ and $P_2$ are nearly identical with $P_3$.




Sketch (l).- Error in $v_3$ using Treanor's modified method (B14).

combined with the step-size control (B12), and started with $\Delta s = 0.005$, this method at once chose a step size equal to 0.01 for which it is marginally stable. The eigenvalue -500 was detected and -1000 excluded as long as this step size was used, see sketch (k). Eventually, however, the step was doubled to the unstable size 0.01. In a few steps the eigenvalue -1000 emerged (as indicated by sketch (k)), and the error phenomenon typical of this kind of step control ensued.

Certainly the curves for $\vec{w}$ in sketches (d), (g), and (j) could be made much smoother with more refined techniques for modifying the step size. However, two things must be borne in mind: First, such refinements reduce the total distance the methods advance a solution in a given number of steps; second, the implicit method requires no such refinements to improve the smoothness.

## APPENDIX C

### STABILITY CRITERION FOR TREANOR'S METHOD APPLIED TO $v' = \lambda v$

Treanor's method can be written in predictor-corrector form as

$$v^{(1)}_{n+(1/2)} = v_n + \frac{1}{2} h v_n' \tag{C1a}$$

$$v^{(2)}_{n+(1/2)} = v_n + \frac{1}{2} h v^{(1)'}_{n+(1/2)} \tag{C1b}$$

$$v^{(3)}_{n+1} = v_n + h \left[ 2v^{(2)'}_{n+(1/2)} F_2 + v^{(1)'}_{n+(1/2)} PhF_2 + v_n'(F_1 - 2F_2) \right] \tag{C1c}$$

$$\begin{aligned} v_{n+1} = v_n + hv_n'F_1 + h\delta_3(Pv_n + v_n') + h\delta_2 \left[ Pv^{(1)}_{n+(1/2)} + v^{(1)'}_{n+(1/2)} \right] \\ + h\delta_2 \left[ Pv^{(2)}_{n+(1/2)} + v^{(2)'}_{n+(1/2)} \right] + h\delta_1 \left[ Pv^{(3)}_{n+1} + v^{(3)'}_{n+1} \right] \end{aligned} \tag{C1d}$$

where

$$F_1 = \frac{e^{-Ph} - 1}{-Ph}, \qquad F_2 = \frac{e^{-Ph} - 1 + Ph}{(Ph)^2}, \qquad F_3 = \frac{e^{-Ph} - 1 + Ph - \frac{1}{2}(Ph)^2}{-(Ph)^3} \tag{C2}$$

and

$$\left.\begin{aligned} \delta_1 &= -F_2 + 4F_3 \\ \delta_2 &= 2(F_2 - 2F_3) \\ \delta_3 &= 4F_3 - 3F_2 \end{aligned}\right\} \tag{C3}$$

The stability of equations (C1) is determined by applying them to $v' = \lambda v$. The resulting matrix equation

$$\begin{bmatrix} E^{1/2} & 0 & 0 & -\left(1+\frac{1}{2}\lambda h\right) \\ -\frac{1}{2}\lambda h E^{1/2} & E^{1/2} & 0 & -1 \\ -\lambda P h^2 F_2 E^{1/2} & -2\lambda h F_2 E^{1/2} & E & -1-\lambda h(F_1-2F_2) \\ -(P+\lambda)h\delta_2 E^{1/2} & -(P+\lambda)h\delta_2 E^{1/2} & -(P+\lambda)h\delta_1 E & E-(1+\lambda h F_1)-(P+\lambda)h\delta_3 \end{bmatrix} \begin{bmatrix} v_n^{(1)} \\ v_n^{(2)} \\ v_n^{(3)} \\ v_n \end{bmatrix} = 0 \tag{C4}$$

where $E \equiv \exp[h(d/dx)]$ has a characteristic equation with only one nonzero root. It is given by the equation

$$E = 1 + \lambda h F_1 + \lambda(\lambda + P)h^2[2F_2 - 4F_3 + (4F_3 - F_2)(F_1 + PhF_2)]$$

$$+ \frac{1}{4}\lambda^2(\lambda + P)h^3[2F_2 - 4F_3 + 2F_2(4F_3 - F_2)(2 + Ph)]$$

$$+ \frac{1}{2}\lambda^3(\lambda + P)h^4F_2(4F_3 - F_2) \tag{C5}$$

The combinations of real positive $Ph$ and negative $\lambda h$ which make the right-hand side of equation (C5) equal to unity form the stability boundaries shown in figure 1. One can easily show

$$\lim_{Ph\to\infty} \quad PhF_1 = 1 , \qquad PhF_2 = 1 , \qquad PhF_3 = 1/2$$

Thus, for large values of $Ph$, equation (C5) reduces to

$$E = 1 + \lambda h + \frac{1}{2}\lambda^2 h^2$$

which is the characteristic equation for the second-order, Runge-Kutta method. This gives the asymptotic value of the left stability boundary shown in figure 1.

## REFERENCES


1. Curtiss, C. F.; and Hirschfelder, J. O.: Integration of Stiff Equations. Proc. Natl. Acad. Sci., U.S.A., vol. 38, 1952, pp. 235-243.

2. Emanuel, G.: Problems Underlying the Numerical Integration of the Chemical and Vibrational Rate Equations in a Near Equilibrium Flow. Rep. AEDC-TDR-63-82, Arnold Engineering Development Center, Tullahoma, Tenn., 1963.

3. Treanor, Charles E.: A Method for the Numerical Integration of Coupled First-Order Differential Equations With Greatly Different Time Constants. Math. Comp., vol. 20, no. 93, Jan. 1966, pp. 39-45.

4. Rush, D. G.; and Pritchard, H. O.: Vibrational Disequilibrium in Chemical Reactions. Proc. Eleventh Symposium (International) on Combustion, Berkeley, Calif., Aug. 7-12, 1966.

5. Bray, K. N. C.; and Pratt, N. H.: Conditions for Significant Gasdynamically Induced Vibration-Recombination Coupling. Proc. Eleventh Symposium (International) on Combustion, Berkeley, Calif., Aug. 7-12, 1966.

6. Moretti, Gino: A New Technique for the Numerical Analysis of Nonequilibrium Flows. AIAA J., vol. 3, no. 2, Feb. 1965, pp. 223-229.

7. DeGroat, James J.; and Abbett, Michael J.: A Computation of One-Dimensional Combustion of Methane. AIAA J., vol. 3, no. 2, Feb. 1965, pp. 381-383.

8. Tyson, T. J.: An Implicit Integration Method for Chemical Kinetics. TRW Space Technology Lab. Rep. 9840-6002-RU000, Sept. 1964.

9. Faddeeva, V. N.: Computational Methods of Linear Algebra. (Translated by C. D. Benster), Dover Pub., Inc., 1959.

10. Lomax, Harvard: An Operational Unification of Finite Difference Methods for the Numerical Integration of Ordinary Differential Equations. NASA TR R-262, 1967.

11. Certaine, J.: The Solution of Ordinary Differential Equations With Large Time Constants. Mathematical Methods for Digital Computers, A. Ralston and H. S. Wilf, eds., Wiley and Sons, Inc., 1960, pp. 128-132.

12. Varga, R. S.: Matrix Iterative Analysis. Prentice-Hall, Inc., Englewood Cliffs, N. J., 1962.

13. Flügge-Lotz, I.; and Davis, R. T.: Laminar Compressible Flow Past Axisymmetric Blunt Bodies. Tech. Rep. 143, Div. of Eng. Mech., Stanford Univ., Feb. 1964.

14. Milne, W. E.: Numerical Calculus. Princeton University Press, 1949.

15. Marrone, P. V.: Inviscid, Nonequilibrium Flow Behind Bow and Normal Shock Waves. Parts I and II. General Analysis and Numerical Examples. Cornell Aero. Lab. Rep. CAL-QM-1626-A-12(I and II), 1963.

16. Hall, J. G.; Eschenroeder, A. Q.; and Marrone, P. V.: Inviscid Hypersonic Airflows With Coupled Nonequilibrium Processes. Cornell Aero. Lab. Rep. AF-1413-A-2, May 1962.

17. Parlett, B.: Applications of Laguerre's Method to the Matrix Eigenvalue Problem. Tech. Rep. 21, Applied Math. and Stat. Lab., Stanford Univ., 1962.



NASA-Langley, 1967 —— 19 A-2649